THE UNIVERSITY OF ALBERTA

PSYCHOPHYSICAL RELATIONS OF JUDGED SIMILARITY AND DIFFERENCE

by

© ROBERT PAUL MARKLEY

A THESIS
SUBMITTED TO THE FACULTY OF GRADUATE STUDIES
IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE DEGREE
OF DOCTOR OF PHILOSOPHY

DEPARTMENT OF PSYCHOLOGY

EDMONTON, ALBERTA

FALL, 1970

Abstract

This study investigated the relationship between combinatorial judgments and a physical measure of objectively simple stimuli as effected by the response procedure subjects were instructed to use.

Judgments of the magnitude of similarity and differences between pairs of stimuli are frequently used in several areas of behavioral science. Systematic studies of the differential effect of response procedure on single stimulus magnitude judgments have long occupied psychophysicists. This paper extended this concern to judgments of paired stimuli. Recently developed nonmetric scaling procedures were used in an effort to distinguish between processes operating in perception of the stimuli (inputs) and overt responses or judgments (outputs).

Three groups of subjects made similarity judgments of all possible pairs of nine Munsell neutral greys. Two additional groups made difference judgments of the same pairs. The response procedures were Magnitude Estimation, Category Rating, and Similarity Estimation for the three similarity groups. Difference judgments were made using Magnitude Estimation and Category Rating procedures. Two further groups made Category Rating and Magnitude Estimation judgments of either the lightness or darkness of single grey stimuli.

Abstract

This study investigated the relationship between combinatorial judgments and a physical measure of objectively simple stimuli as affected by the response procedure subjects were instructed to use.

Judgments of the magnitude of similarity and differences between pairs of stimuli are frequently used in several areas of behavioral science. Systematic studies of the differential effect of response procedure on single stimulus magnitude judgments have long occupied psychophysicists. This paper extended this concern to judgments of paired stimuli. Recently developed nonmetric scaling procedures were used in an effort to distinguish between processes operating in perception of the stimuli (inputs) and overt responses or judgments (outputs).

Three groups of subjects made similarity judgments of all possible pairs of nine Munsell neutral greys. Two additional groups made difference judgments of the same pairs. The response procedures were Magnitude Estimation, Category Rating, and Similarity Estimation for the three similarity groups. Difference judgments were made using Magnitude Estimation and Category Rating procedures. Two further groups made Category Rating and Magnitude Estimation judgments of either the lightness or darkness of single grey stimuli.

Nonmetric scaling analyses indicated that the greys were perceived to vary on a one dimensional continuum. Scale values (location) on the dimension represented a power transformation of a physical measure of reflectance.

The nonmetric analyses also indicated that overt judgments of the difference between greys were related by a power function to differences on the subjective scale. These findings supported a two stage model of magnitude judgment first suggested by Attneave (1962) and developed by Curtis, Attneave, and Harrington (1968). The model was

$$J = a(\emptyset_i^{\,k} - \emptyset_j^{\,k})^m + b,$$

where $i>j$, J is a magnitude judgment, $\emptyset$ a measure of reflectance, and a, k, m, and b are parameters estimated from the data.

A model was proposed to relate similarities to differences on the subjective continuum. This model was $S = G - D$ where S is a subjective similarity, G a constant, possibly related to the maximum range of stimulation, and D is a function of $\emptyset_i^{\,k} - \emptyset_j^{\,k}$. Judgments from two of the similarity groups were found to fit the model. The Magnitude Estimations of Similarity were not described well. Alternative models were examined. These also failed to convincingly describe the Magnitude Estimation of Similarity results.

TABLE OF CONTENTS

TABLE OF CONTENTS (cont.)

LIST OF TABLES

LIST OF FIGURES

LIST OF APPENDICES

INTRODUCTION

Similarity judgments have proved to be a valuable tool in studies of perception and cognition. Shepard (1966) has described the similarity judgment as a general measure of the substitutability or equivalence of stimuli. Such judgments contain important information about the ways in which stimuli are perceived or coded and are (a priori) less subject to biases that are introduced when more specific judgmental sets (e.g., loudness or complexity) are used. The manner in which stimuli are equivalent can be determined by analyses of similarity judgments. The importance of particular stimulus attributes or elements is not prejudged by the scientist nor are these attributes allowed to predetermine the nature of the analyses (Hake & Rodwan, 1966).

Analyses of similarity data usually begin by viewing the stimulus objects as points in a multidimensional metric space. The magnitudes of the similarity judgments are assumed to reflect in some way the magnitudes of the interpoint distances in the space. The scaling problem is then to attempt to determine the dimensionality of the space, the spatial coordinates of the stimuli and a function relating judgments (responses) to the interpoint distances.

One alternative view was proposed by Ekman (1963). He

suggested that similarity be interpreted as a scalar product of two vectors. Very little has been done with this model outside of Sweden. Most investigators have preferred a distance model. The distance model has been found to provide a more satisfactory account of some of Ekman's earlier data (cf., Shepard, 1962).

Until recently analytic techniques have required that the data consist of either scalar products (Ekman, 1963) or ratio scale measures of distance (cf., Torgerson, 1958). Obtaining these from rather crude similarity responses was usually either not practical or a theoretically suspect procedure. The scaling analysis of similarity judgments became a feasible technique with the development of various nonmetric analyses (cf., Shepard, 1962; Kruskal, 1964; Torgerson, 1965; McGee, 1966). These procedures use only the rank order of similarity (or dissimilarity) measures.

Similarity measures between stimuli, in addition to their usefulness in attempting to scale stimuli, can also be used directly as independent variables or covariates in studies of learning and cognitive processes. Runquist (1966) discusses this tactic. Melton and Safier (1951) provide a set of adjective pairs, scaled for similarity, for use in paired associate learning tasks.

Scale Invariance

There are as many ways to obtain similarity judgments as there are psychophysical scaling procedures. If similarity judgments are to be used as a scale of interpoint distance then different procedures that lead to different scales would generate empirical and theoretical paradoxes. The same problem could arise, during the more common task of using similarity judgments to scale stimuli by locating them in a multidimensional space, if different scales of similarity produced different configurations.

Scaling literature abounds with conflicts over the lack of invariance between scaling methods. Galanter (1962) and the two Annual Review papers by Ekman and Sjoberg (1965) and Zinnes (1968) amply chronicle this issue. The problem, in essence, is that magnitude estimation and its related procedures will, in the great majority of experiments, produce a judgmental function that is not linearly related to the function produced by the various categorical procedures. However, there are certain qualitative "metathetic" continua (e.g., pitch, angle of inclination, and proportionality) for which it is claimed that the different procedures produce invariant judgmental functions (Stevens & Galanter, 1957). But, there are

doubts about the existence of a metathetic-prothetic distinction (see Warren & Warren, 1963).

There are at least two studies that suggest that perceived similarity may have properties of metathetic continua. Ekman and Waern (1959) reported that a form of category rating and ratio estimation judgments of the similarity of pairs of circles were linearly related. They concluded from this that the category judgments of similarity formed a ratio scale. Markley, Ayers, and Rule (1969), using independent groups, reported that Similarity Estimations (0-100 similarity judgments) were linearly related to Magnitude Estimations of Similarity. The stimuli were lines differing in length. A one dimensional nonmetric scaling solution for the data was logarithmically related to physical length.

Of course, if only the newer nonmetric procedures were to be used, then concern over invariance of similarity judgments is unimportant so long as the various measures were monotonic to one another.

Similarities in One Dimension

There has also been some recent interest in use of similarity judgments to develop unidimensional psychological scales. A few years ago Ekman and associates at

the University of Stockholm (Ekman & Waern, 1959; Eisler & Ekman, 1959) studied similarity judgments of physical stimuli that were varied in only one dimension--pitch, circle size, heaviness, and darkness of greys. Ekman, Goude, and Waern (1961) reported that a similarity judgment (S_{ij}) for two stimuli could be predicted from ratio estimation (R_i) scale values by the function $S_{ij} = 2R_i/(R_i + R_j)$ when $i < j$. At the time, it was thought that ratio estimation scale values were context free.

An important point, apparently missed by Ekman and co-workers, was made by non-psychologists Carmichael, Julius, and Martin (1965). These authors pointed out that similarities of stimuli varying in one dimension should be relative to the maximum difference (context) expected by the <u>S</u>. Torgerson (1965) reported empirical findings supporting this notion for a class of multidimensional stimuli.

<u>Similarities and Differences</u>

The similarity of two stimuli is often thought of as an indirect indicator of the distance between stimuli in a psychological space. Several authors have used the terms difference, distance, and similarity interchangeably as measures of psychological proximity with no regard for possible biases in outcomes due to <u>S</u>s' differing

interpretations of the terms. For example, Attneave (1950) instructed his Ss to respond "on a basis of overall similarity" yet used rating scales with extreme categories labelled "identical" and "extremely different" (p. 523). Other researchers have used physical distance measures obtained by asking Ss to place a color chip on a grid in such a way that grid distance was matched to subjective difference or similarity (Helm, 1964; Indow & Uchizono, 1960; Indow & Kanazawa, 1960). Once again, if only ordinal information is required there is no problem as long as there is a monotonic relation between the various measures. But if a higher scale type is required then concern over the exact set imparted to Ss by the scaling instructions is warranted.

A magnitude estimation of difference judgment has been used with grey stimuli by Torgerson (1961) to study Ss' use of numerals as a response procedure. Difference judgments have more recently been used by Curtis, Attneave, and Harrington (1968) and by Fagot and Stewart (1969) to test sequential models of magnitude judgment. As far as is known the empirical comparison between similarity and difference judgments suggested here is original. At the theoretical level, one author (Landahl, 1945) has

suggested that difference judgments and similarity judgments represent different psychological processes. Landahl has described two types of theoretical neural net mechanisms capable of mediating each.

Response Factors

In addition to the problems of scaling procedure outlined above, the present study provides an opening to one of the major substantive problems of contemporary psychophysics. A shortcoming of recent psychophysical scaling theory has been the absence of models of magnitude judgment that incorporate separable response and sensory parameters. Stevens' (1957) power function has had no real challenger, although a significant portion of recent research has been devoted to the demonstration of the effect of nonstimulus variables on the power law's exponent (a recent comprehensive survey was done by Poulton, 1968).

The problem of response variables was pointed out as far back as 1956 by Garner, Hake and Eriksen (1956). Little has been done to integrate response biases into theory. Rule and co-workers (Rule, 1966, 1968, 1960; Markley, 1965; Rule & Markley, 1970) have used correlational procedures to demonstrate the existence of response biases and to assess their effect on Ss' judgments. However, demonstration

that an effect exists does not necessarily fit it into a theoretical description of the behaviors in question.

The present research was originally designed to utilize the properties of Kruskal's nonmetric scaling analysis which produces a spatial configuration of stimulus points and the relationship between Ss' responses and interpoint distances to separately provide a description of effects due to sensory and response systems underlying the similarity and difference judgments. Subsequent to the collection of the present data Curtis, Attneave, and Harrington (1968) reported a test of a two-stage model of magnitude judgment originally proposed by Attneave (1962).

Two Stage Model

The development of the two stage model (Attneave, 1962; Curtis, Attneave, & Harrington, 1968) rests on two assumptions: (1) that the power law was valid for psychological magnitudes (Ψ); i.e.,

$$\Psi_i = a_1 \phi_i^k, \tag{1}$$

where ϕ_i is the physical stimulus magnitude; (2) that a numerical judgment (J) was a power function of psychological magnitude; i.e.,

$$J_i = a_2 \Psi_i^m. \tag{2}$$

Then, the usual magnitude estimation scaling operation

would produce a function

$$J_i = a'\emptyset_i^{km}, \quad (3)$$

where k and m are input and output parameters, respectively. Data from the usual magnitude estimation procedures do not provide separate estimates of k and m, and results are simply described by

$$J_i = a\emptyset_i^n , \quad (4)$$

where n = km.

Curtis _et al_. (1968) suggested that if _S_s were required to make magnitude judgments of the difference between unidimensional stimuli, then the input and output operations should produce data described by the equation

$$J_{ij} = a(\emptyset_i^k - \emptyset_j^k)^m , \ (i>j). \quad (5)$$

An additive constant is often included to correct for small curvatures at the extreme values of psychophysical data. Inclusion of the constant yields

$$J_{ij} = a(\emptyset_i^k - \emptyset_j^k)^m + b, \ (i>j). \quad (6)$$

Thus, if S_s make magnitude judgments of single stimuli and judgments of the differences between stimuli, it should be possible to obtain separate estimates of n (Eq. 4) and m and k (Eq. 6). Support for the model is obtained if Equations 4 and 6 do adequately describe the data and n=km. This was the finding of Curtis _et al_. (1968) for magnitude

judgments of weights. Their functions based on group data were

$$J = 1.722\ (\phi)^{.746} - 4.852,$$

and

$$J_{ij} = 2.252\ (\phi_i{}^{.645} - \phi_j{}^{.645})^{1.141} - .671.$$

The value of m x k (.645 x 1.141) was .736 which is quite close to the independently obtained value of n, .746. This finding was confirmed by separate analyses of each of the ten Ss in the Curtis et al. (1968) study. Curtis also reports confirmation of the model for summation judgments (Curtis & Fox, 1969), and extended the findings to category ratings of differences (Curtis, 1970).

The two stage model suggests that a change in response procedure should manifest itself in a change of the value of m, the output exponent. The input exponent k should remain constant. The present study should provide further information about this model's ability to describe the judgment process when response procedures are changed and when there is a change in the relationship judged (i.e., difference vs. similarity).

The present study attempted to provide further information about the judgment of similarity. To do this the interrelationships of familiar psychophysical scaling

procedures used in obtaining measures of judged similarity were investigated. The effects of response procedure on judgments of the difference between stimuli and the relationship between judged differences and judged similarities were also observed. Since the purposes of the study were primarily to look at effects of response procedure on a complex judgment, relatively simple stimuli that varied on one physical dimension were chosen. Individual differences in cognitive or perceptual processing of the stimuli should be minimized with stimuli of this type. It is well known that individual Ss display different rules for combining information from several dimensions (e.g., Cliff, 1968, and Shepard, 1964).

METHOD

Stimuli. The stimuli were nine Munsell grey papers: N3.5, N5.0, N6.0, N7.0, N7.5, N8.0, N8.5, N9.0, N9.5. These were chosen so that reflectance would vary in approximately equal steps. Actual reflectance was measured by a Gamma Scientific Inc. Model 700 log-linear photometer. The photometer was calibrated against the Gamma No. 220 Standard Lamp Source and Gamma No. 700-15 Standard of Reflectance (92%). The measured reflectance values averaged over several days and lighting conditions were 9.5, 20, 30.5, 43, 51, 58, 69, 80, and 87 percent. The median deviation from reflectance values supplied by the Munsell Corp. was 1.6%. All deviations are within the limits of error of the Model 700 photometer.

The stimuli were 4.8 cm square patches cut from the grey papers and mounted either singly or in paris on 25.4 cm square white poster paper (average reflectance = 98%). There were nine single stimulus cards and 36 cards with paired stimuli.

Subjects. All Ss were paid volunteer students at The University of Alberta 1967 Summer Session. Each S participated in a single experimental session lasting up to

45 minutes.

Procedure. All experimental sessions were held in a research room at the University of Alberta. The room was bisected by a black screen (reflectance = 06%). There was a 76.2 x 78.7 cm opening located at table height in the screen. The stimulus cards were presented on a small black tilted (35^{o}) podium placed in the center of the opening. Black (05%) curtains hung behind the stimulus tray prevented _S_s from seeing the experimenter and the back half of the room. Illuminance incident to the stimuli was approximately 645 lux. The _S_s were seated at a 53 x 61 cm table. The front edge of the table was 76 cm from the base of the stimulus tray. The _S_s recorded their responses on data sheets placed on the table.

In all conditions the _E_ stood behind the screen and presented stimuli by placing a stack of stimulus cards (a blank card showing) on the podium. When a _S_ was ready, the blank was removed, revealing the first stimulus to be judged.

Conditions. _S_s were randomly assigned to one of seven scaling tasks: (1) Similarity Estimation, (2) Magnitude Estimation of Similarity, (3) Magnitude Estimation of Differences, (4) Category Rating of Similarity

(5) Category Rating of Differences, (6) Lightness Scaling, and (7) Darkness Scaling.

The Similarity Estimation group judged the similarity of pairs of greys on a scale ranging from zero (no similarity) to 100 (identical). The 36 pairs were presented twice to each S in a pseudo-random order. Several orders were used, and potential spatial errors were counterbalanced.

The Magnitude Estimation of Similarity group judged the similarity of pairs using the method of magnitude estimation. A stimulus pair (N7.0-N8.5) was designated as a standard, and its similarity was assigned the value 10. The standard pair was presented and identified by E once at the beginning and once half way through a session. The Ss judged each pair twice. The instructions were typical of magnitude estimation tasks.

The Magnitude Estimation of Difference Ss were instructed to judge the difference between the stimuli using the magnitude estimation technique. The standard pair and procedure were the same as they were in the previous Magnitude Estimation group except that the Ss were instructed to attend to the differences between the stimuli. Attributes contributing to stimulus differences were not specified.

The two category ratings groups rated either the similarity or differences of the paired greys on a seven point category scale. Two standard pairs were shown at the beginning and at the midway point of the session. Pair N3.5-N9.5 was an example of a _highly different_ (_least similar_) pair, and N7.0-N7.5 served as an example of a pair that was _least different_ (_highly similar_).

Subjects in the two groups which received only single stimuli made both category ratings and magnitude estimations of either the lightness or darkness of single grey patches. For magnitude estimation a standard grey (N7.5) was assigned a modulus of 10. N9.5 (Very Light or Not At All Dark) and N3.5 (Very Dark or Not At All Light) were standards for the category ratings. Each _S_ judged each stimulus four times with each response procedure.

Although no _S_ who was asked to judge lightness was asked to judge darkness (and vice versa), several _S_s reversed themselves at times during the session. Data from the oscillating _S_s were not included in subsequent analyses. The discarded data, however, were consistent with data collected under the opposite instructional set.

Twelve _S_s were originally included in each group.

One S was excluded from each of the Category Rating groups due to failure to follow instructions concerning the marking of their response sheets. In the two single stimulus scaling conditions, Ss tended to spontaneously reverse their scales of judgments. In each of these groups only seven Ss were able to successfully complete both tasks.

RESULTS

Responses were averaged across Ss and trials for each pair of stimuli by taking geometric means in the Magnitude Estimation conditions, medians in the Category Rating conditions, and arithmetic means for the Similarity Estimations. When scoring category responses, the seven steps of the scale were assigned the integers one through seven with the least light (least dark, least different, or least similar) end of the scale scored as one. A summary of the data from each of the conditions is given in Appendices I and II.

Curve Fitting Procedures

All functions were fitted to a least squares criterion. A computer program computed parameter estimates $A_1, A_2, \ldots, A_r$ by minimizing the expression

$$Q = \sum_{i=1}^{n} W_i [Y_i - f(A_1, A_2, \ldots, A_r; X_{1i}, X_{2i}, \ldots X_{mi})]^2,$$

where W_i represents a weight applied to the *i*th observation, Y_i is the *i*th value of the dependent measure, X_{ij} is the *i*th value of the *j*th independent measure, and *f* denotes the theoretical function. When magnitude estimations were used as the dependent variable the weights, W_i, were set equal

to $1/Y^2$. This weighting takes into account the decreased reliability of large magnitude estimation responses. For other variables the weights were set at 1.0.

Lightness and Darkness Scaling

Results from the two unidimensional scaling conditions, lightness judgments and darkness judgments, are summarized in Appendix II. The psychophysical relations for lightness are illustrated in Figures 1 and 2. The least squares power function for the magnitude estimation of lightness data yielded an exponent of 1.736, slightly higher than exponents reported by Stevens and Galanter (1957), and Mashhour and Hosman (1968), but less than that calculated by Curtis (1968) for the data of Torgerson (1960). The category lightness results of Figure 1 could also be described as a power function of reflectance. The exponent was .64.

Figure 3 shows the relation between magnitude estimations and category ratings as a power function. The constant, .636, was taken from the equation in Figure 1. This was consistent with the results shown in Figures 1 and 2.

Figures 4 and 5 suggest that for the category methods lightness is the reverse of darkness, while for the magnitude estimation technique, lightness is the reciprocal of darkness. The results replicate those of Torgerson (1960).

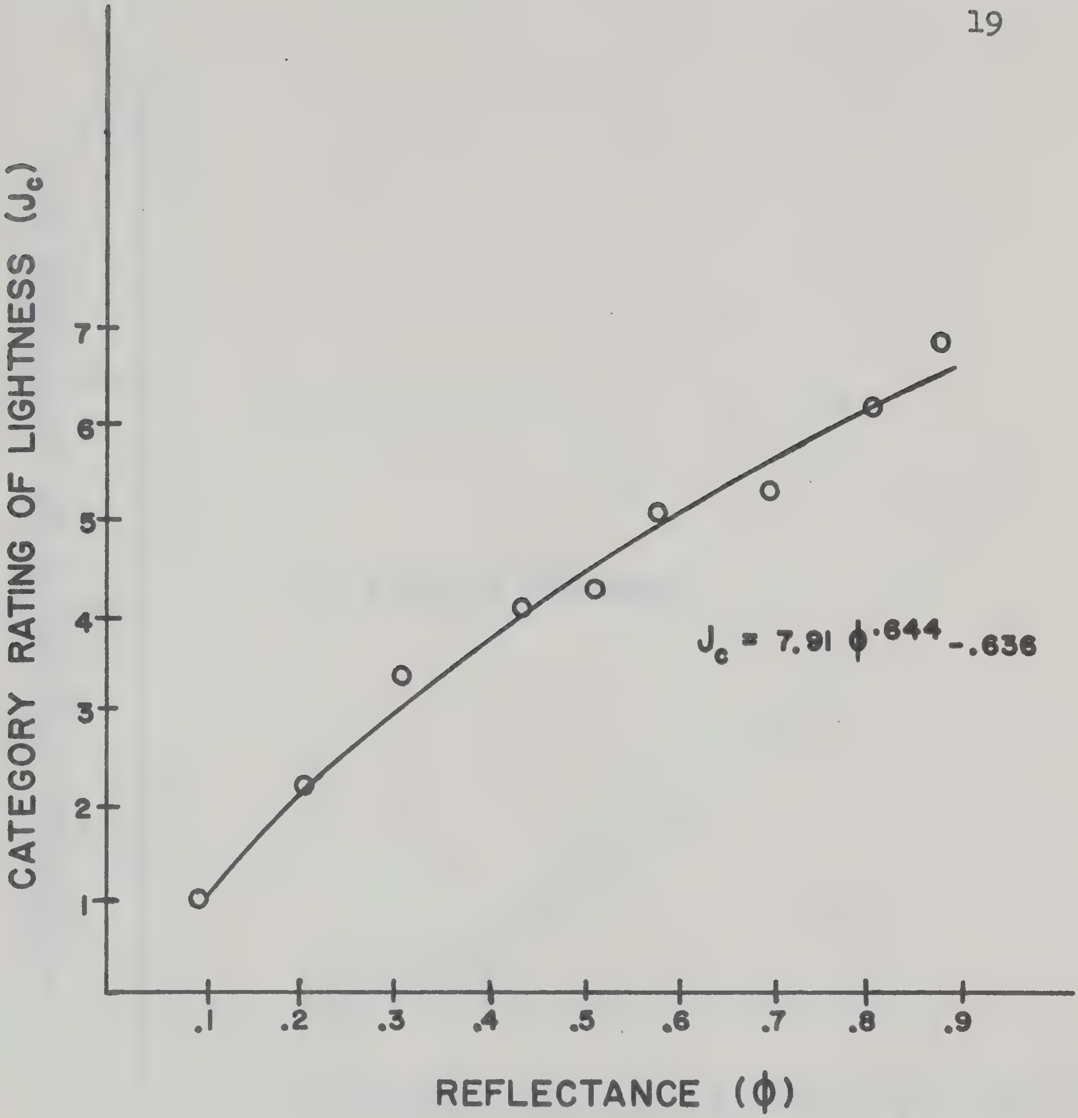

Figure 1. Category ratings of lightness as a function of reflectance.

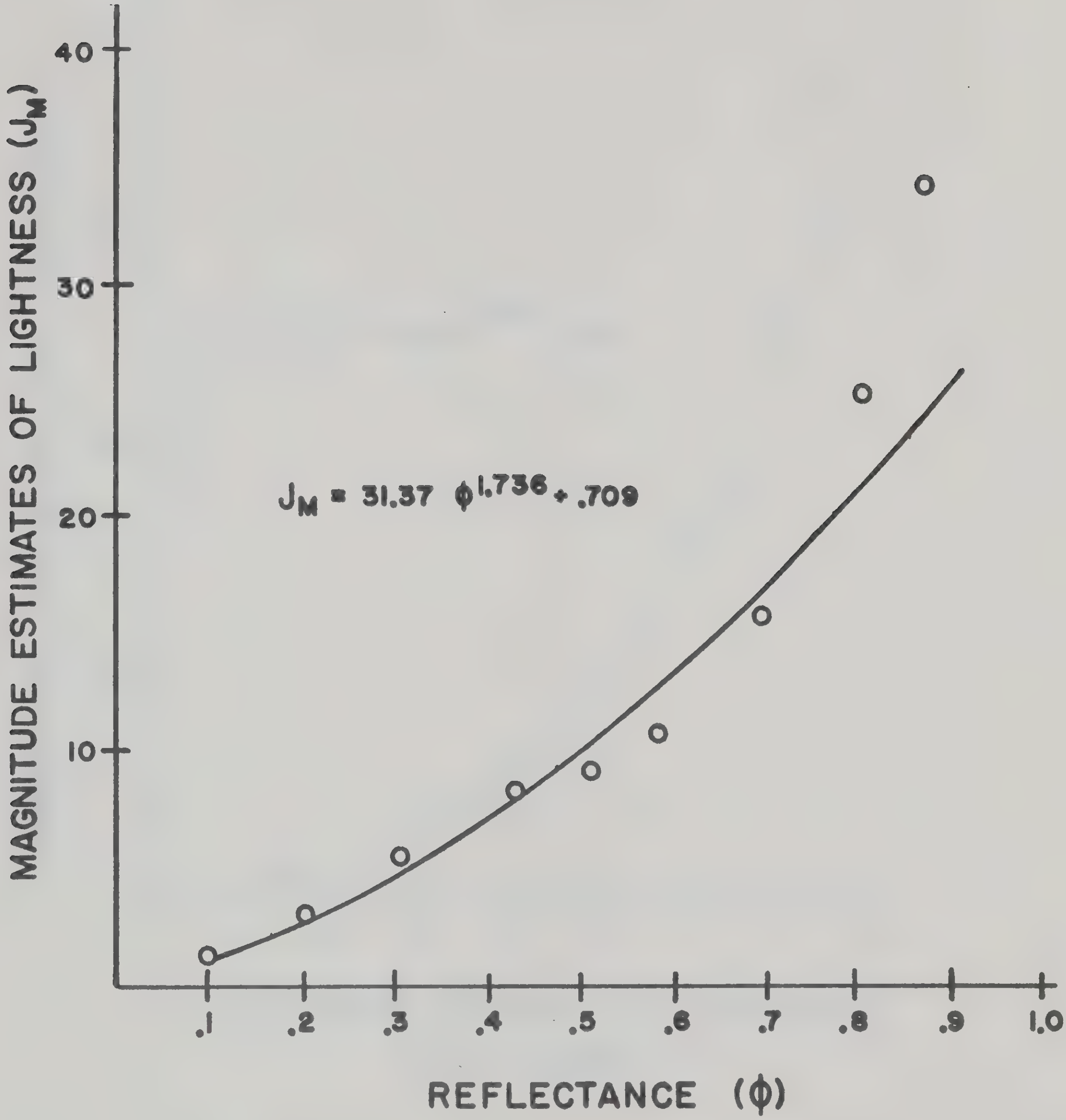

Figure 2. Magnitude estimation of lightness as a function of reflectance.

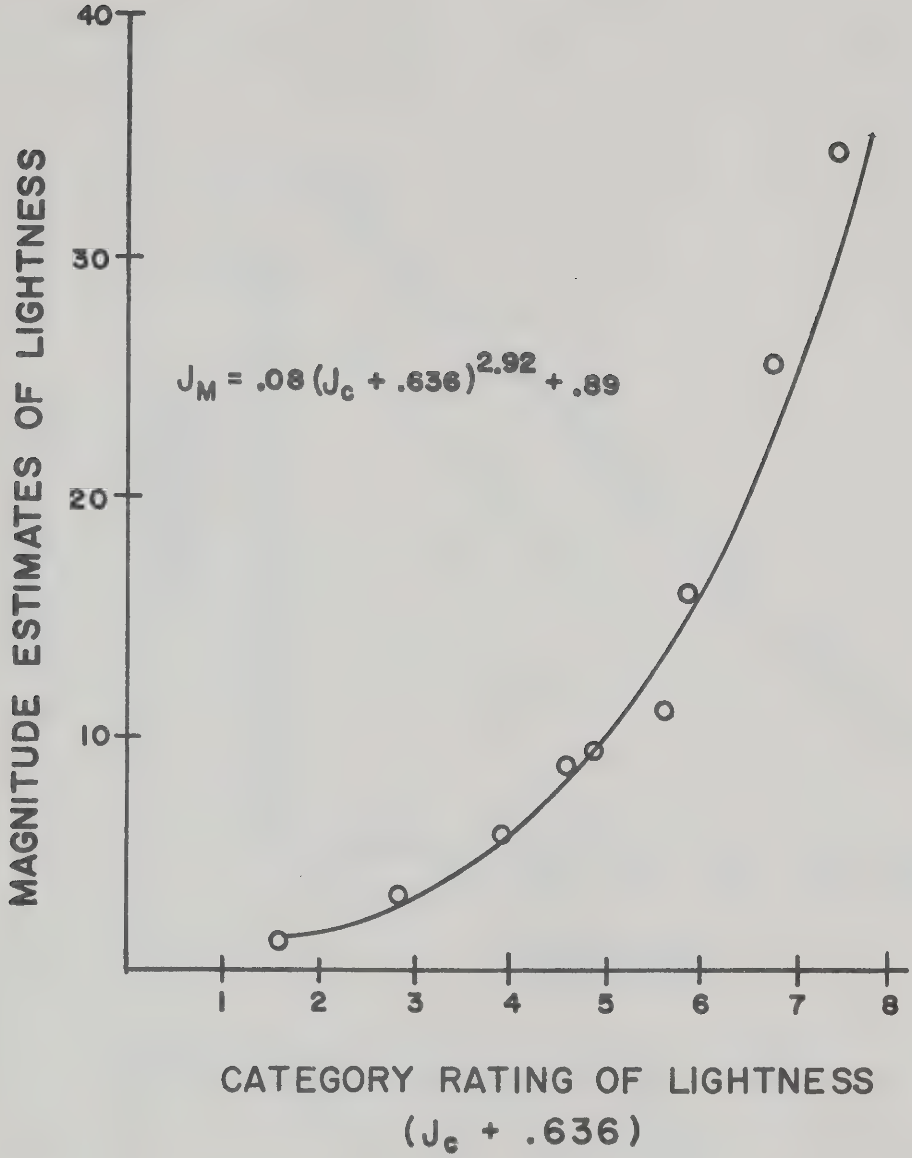

Figure 3. Magnitude estimates of lightness plotted as a function of category ratings of lightness. The constant, .636, added to each category rating was obtained from the least squares solution shown on Figure 1.

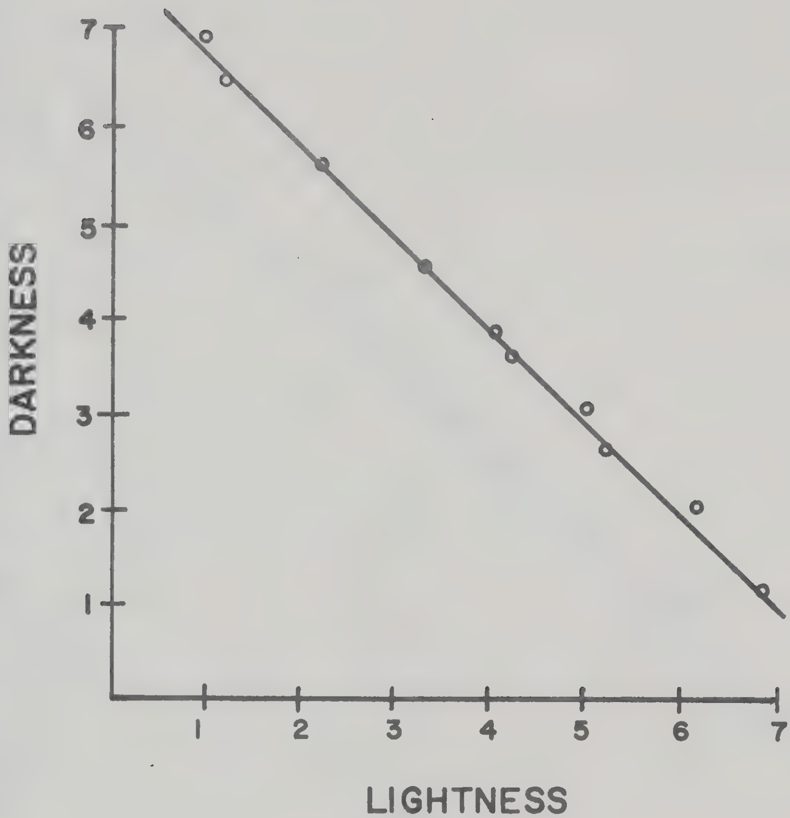

Figure 4. Category ratings of darkness as a function of category ratings of lightness. The straight line was fitted by eye .

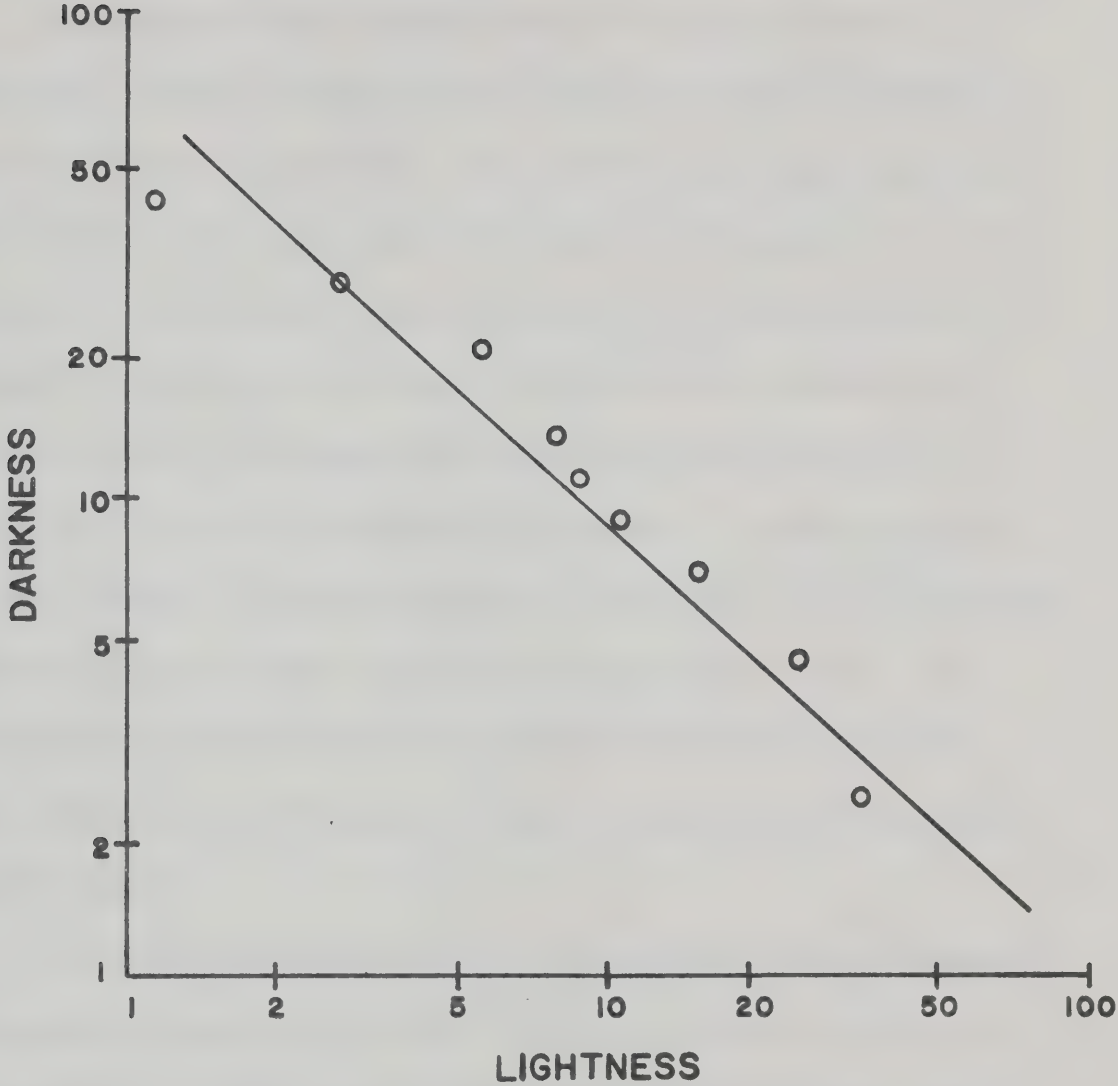

Figure 5. Magnitude estimations of darkness as a function of magnitude estimates of lightness in log-log coordinates. The straight line was fitted by eye.

Multidimensional Analyses

A nonmetric multidimensional scaling procedure (Kruskal, 1964) was applied to the five sets of difference and similarity judgments. The scaling procedure produces a monotonic function relating response measures to a distance in a conceptual space as well as a configuration of stimulus points in the space. The relation between the configuration and known physical variables provides a description of the perceptual or sensory transformation of the stimuli (a psychophysical function). The functional relationships between interpoint distances in the configuration and the several response measures should be indicative of the biasing effects of the various response procedures.

Several configurations (or solutions) were originally obtained for each data set. Different initial configurations and distance metrics were used and the best (least stressful) solution for a given number of dimensions was retained. All results reported here used a Euclidian distance metric. The computer program iterated each configuration until either a goodness of fit measure, stress, fell below .005 or a minimum stress was reached (i.e., the program could find no way to change the configuration so

as to reduce stress). The program was allowed up to 100 iterations to find a satisfactory solution. The primary approach to tied judgments (Kruskal, 1964) was used throughout. The stress statistic (S*) is a measure of the absence of a monotonic relationship between the response measures and the configurational distances. The stress criterion of .005 (an arbitrary "satisfactory good fit") was deliberately chosen to be a smaller than previously published satisfactory stress values, usually .05, (Kruskal, 1964) in order to allow the program full opportunity to find a minimum stress solution.

One dimensional solutions produced stress values well within the satisfactory levels defined by Kruskal (1964). These are shown in Table 1. The scale values reported in Table 1 were highly correlated with each other and clearly ordered the stimuli on the physical dimension of reflectance. The relationship between scale values and reflectance (shown in Figure 6) was nonlinear. The largest deviations from linearity occurred at the low reflectance end of each dimension.

It has been previously reported that scale values obtained from a nonmetric scaling of line length were logarithmically related to physical length (Markley,

TABLE 1

NONMETRIC SCALING ONE DIMENSIONAL SOLUTIONS (SCALE VALUES)

FROM JUDGMENTS OF PAIRS OF GREY STIMULI

STIMULI		DATA SET				
MUNSELL VALUE	REFLECTANCE	CATEGORY RATING SIMILARITY	MAGNITUDE ESTIMATION SIMILARITY	SIMILARITY ESTIMATION	MAGNITUDE ESTIMATION DIFFERENCE	CATEGORY RATING DIFFERENCE
9.5	.87	1.492	1.510	1.450	1.360	1.370
9.0	.80	1.158	1.050	1.090	1.020	1.048
8.5	.69	.702	.720	.786	.710	.745
8.0	.58	.324	.338	.362	.471	.429
7.5	.51	.099	.030	.090	.180	.139
7.0	.43	-.175	-.137	-.240	-.087	-.116
6.0	.305	-.712	-.571	-.639	-.602	-.554
5.0	.20	-1.206	-1.115	-1.167	-1.120	-1.176
3.5	.095	-1.682	-1.834	-1.766	-1.925	-1.885
STRESS (S*)		.026	.044	.024	.035	.030

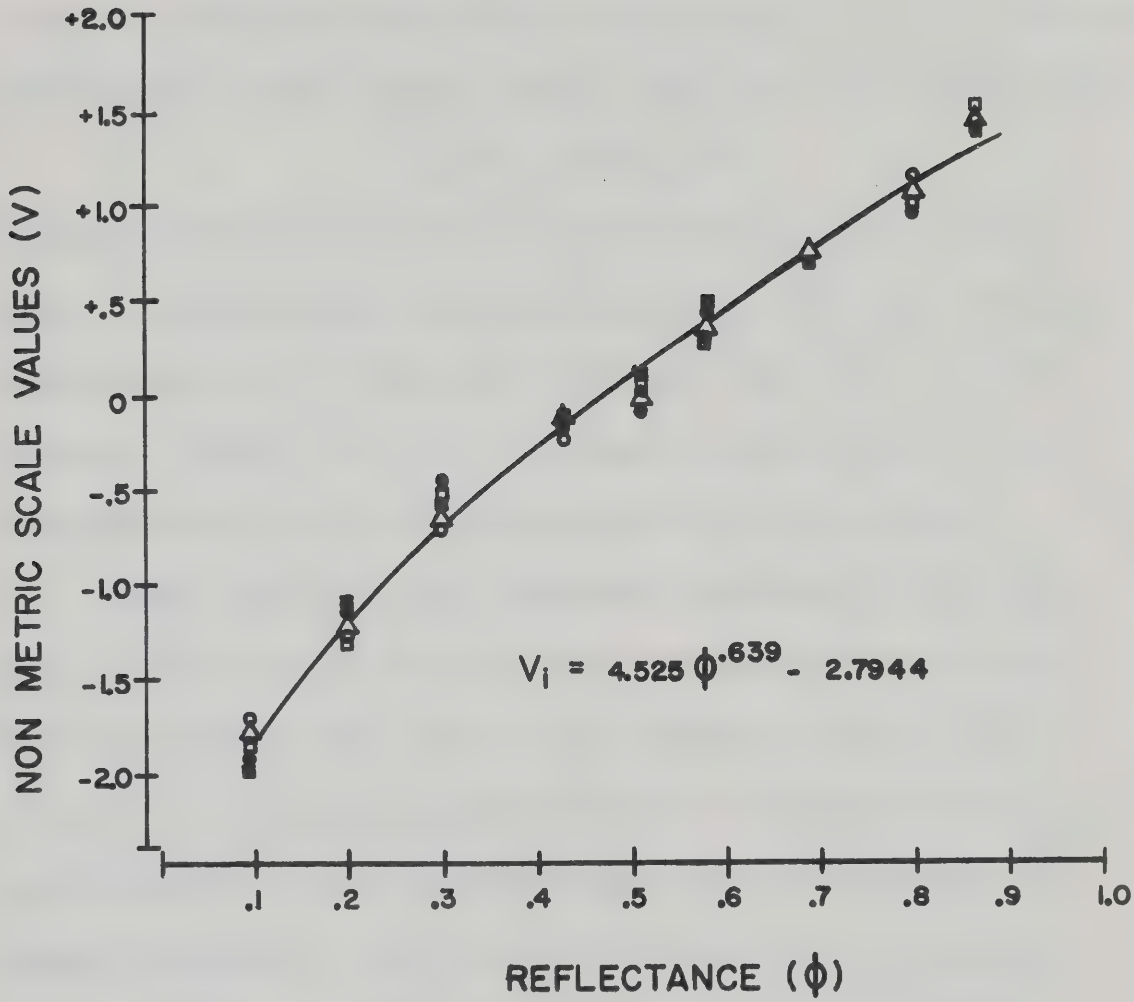

Figure 6. Nonmetric scale values as a function of reflectance. Filled squares are scale values from magnitude estimates of difference. Filled circles are scale values from category ratings of difference. Triangles are scale values derived from similarity estimations. Open circles are derived from similarity ratings. Open squares are scale values from magnitude estimations of similarity. The solid line function is based upon all five sets of scale values.

Ayers, & Rule, 1969). Logarithmic functions relating scale values to physical reflectance did not provide a satisfactory description of the present data. The best log function was

$$V = 1.42 \log_e \emptyset + 6.27, \qquad (7)$$

where V is the scale value, and $\emptyset$ is reflectance of the greys. Within limits of rounding error, the curve fitting program arrived at Equation 7 for all five sets of scale values. Between 92 and 97 percent of the variance of the dependent measure was accounted for by this function.

Power functions with fractional exponents were found to provide a better description (accounting in each case for more than 99 per cent of the variance of the scale values) of the relationships between the scale values and reflectance. The parameters of the power functions are shown in Table 2. Equation 8 is the result of fitting a single function to all of the 45 Scale Values in Table 1

$$V_i = 4.525\ \emptyset^{.639} - 2.794. \qquad (8)$$

Figures 7 and 8 show the category and magnitude estimation of lightness and darkness plotted as a function of the scale values from the nonmetric scaling analyses. The abscissa values were obtained by averaging the scale values from Table 1, then adding the constant 2.794 from Equation 8. Figure 7 indicates that unidimensional category responses

TABLE 2

PARAMETERS OF FUNCTIONS RELATING SCALE VALUES FROM NONMETRIC ONE DIMENSIONAL SOLUTIONS TO REFLECTANCE

$(V = a\emptyset^k + b)$

CONDITION	a	k	b	r^2
Similarity Estimation	4.56	.63	-2.84	.997
Magnitude Est. of Similarity	4.43	.67	-2.68	.997
Category Rating of Similarity	4.166	.86	-2.21	.997
Magnitude Est. of Difference	5.44	.43	-3.85	.997
Category Rating of Difference	5.07	.498	-3.44	.997

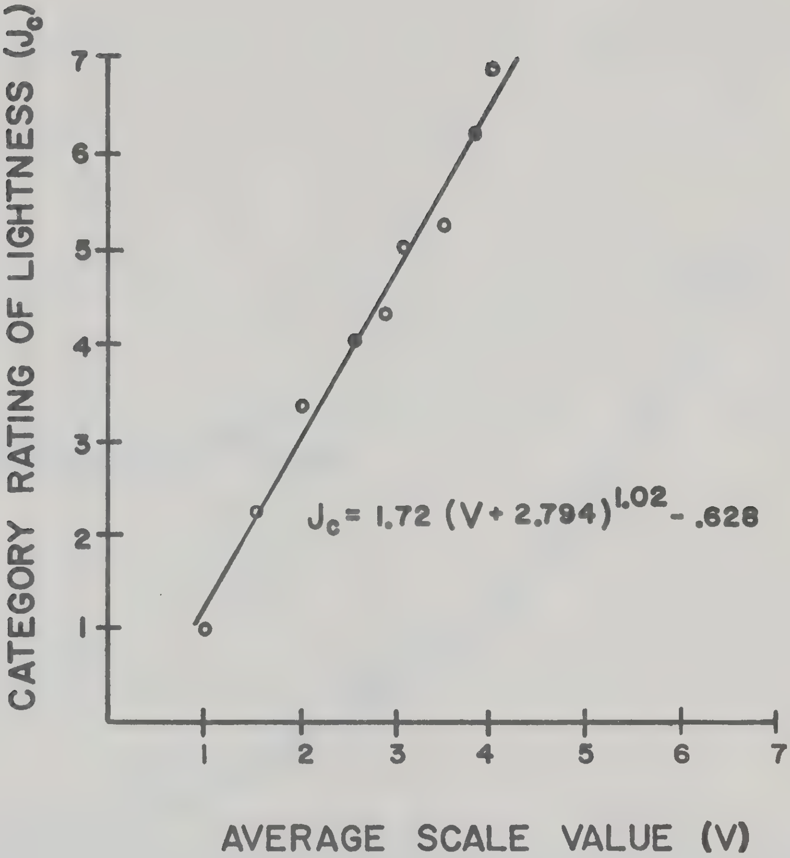

Figure 7. Category ratings of lightness as a function of the average nonmetric scale value of each grey stimulus. The constant, 2.794, added to each scale value is taken from the function shown in Figure 6.

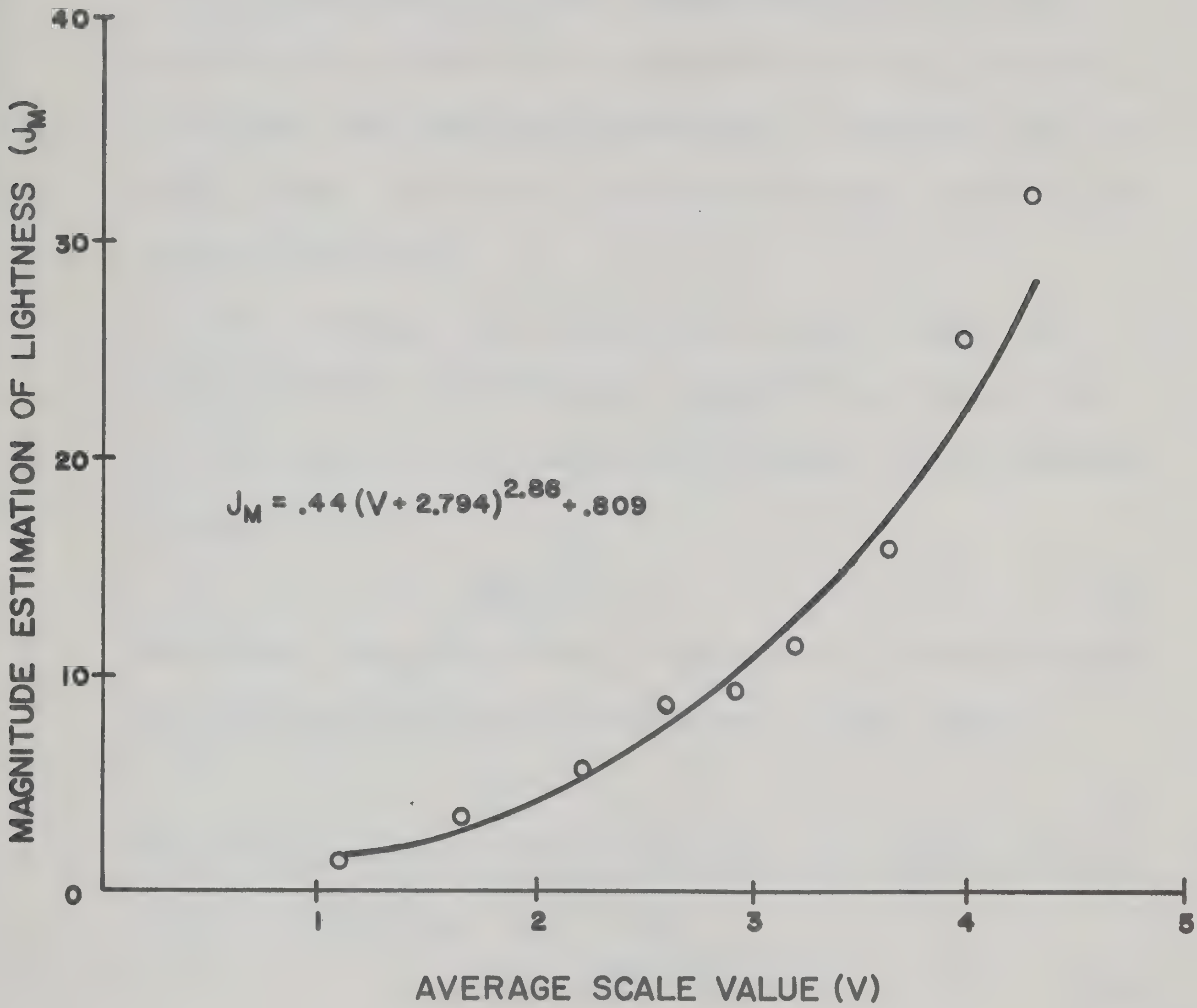

Figure 8. Magnitude estimation of lightness as a function of average nonmetric scale value. The constant added to each scale value is taken from Figure 6.

can be described as a power transformation of the non-metric scale values with an exponent of 1.02. Figure 8 shows that the magnitude estimation of lightness data are also a power function of the nonmetric scale values. The exponent was 2.86.

The finding that the nonmetric scale values are a power function of reflectance provides some support for the two stage model of magnitude judgment. The functions were of the form

$$V_i = a\emptyset_i^k - b. \tag{9}$$

The common observation that judgments are a power function of reflectance was replicated in the present study,

$$J = a_2\emptyset^n + b_2. \tag{10}$$

Rearrange Equation 9 then substitute into Equation 10 and let $m = n/k$ and $a_3 = a_2a^{-m}$ and

$$J = a_3(V_i + b)^m + b_2 \tag{11}$$

as was observed in Figures 7 and 8.

Empirical estimates of Equations 9, 10, are found on Figures 1, 2, and 6. That the input transformation described by Equations 8 and 9 was a power function is consistent with the requirements of the two stage model.

Difference Functions

Figures 9 and 10 show the judged differences plotted

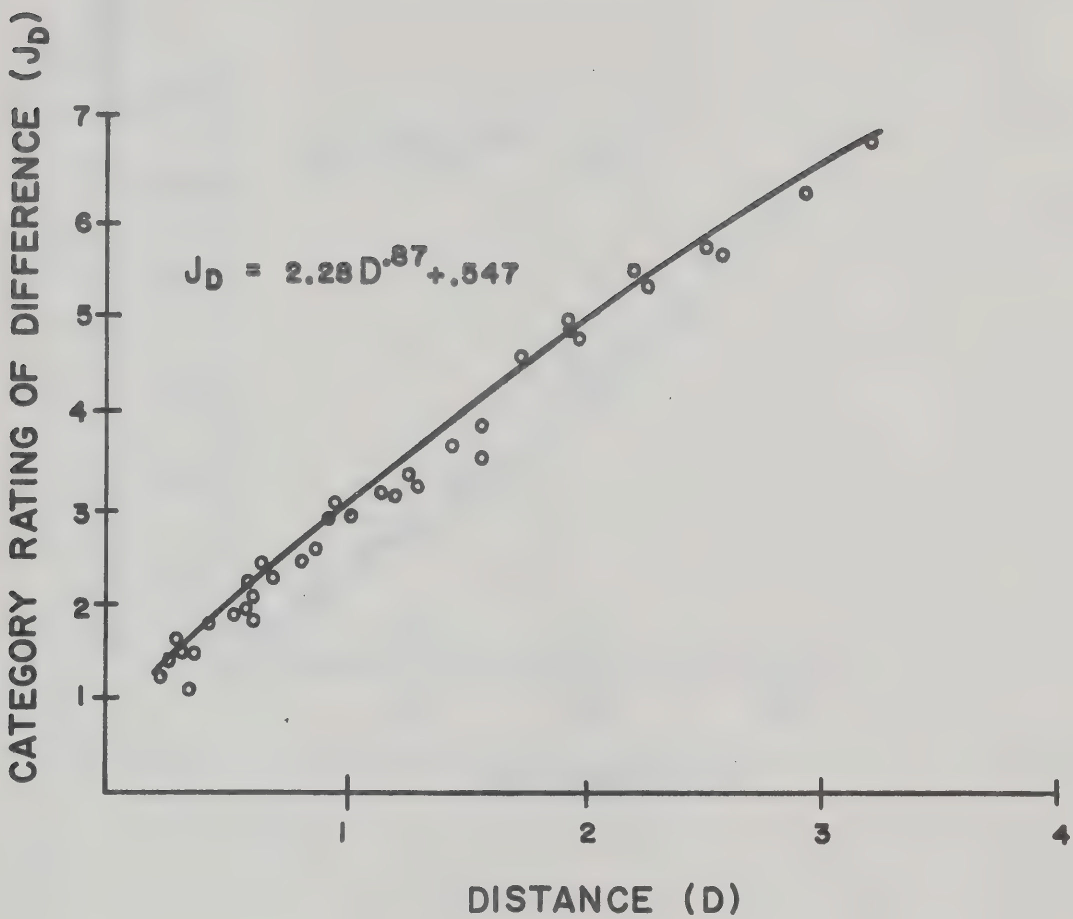

Figure 9. Category ratings of difference as a function of interstimulus distance (D) in the one dimensional space derived from nonmetric scaling analysis.

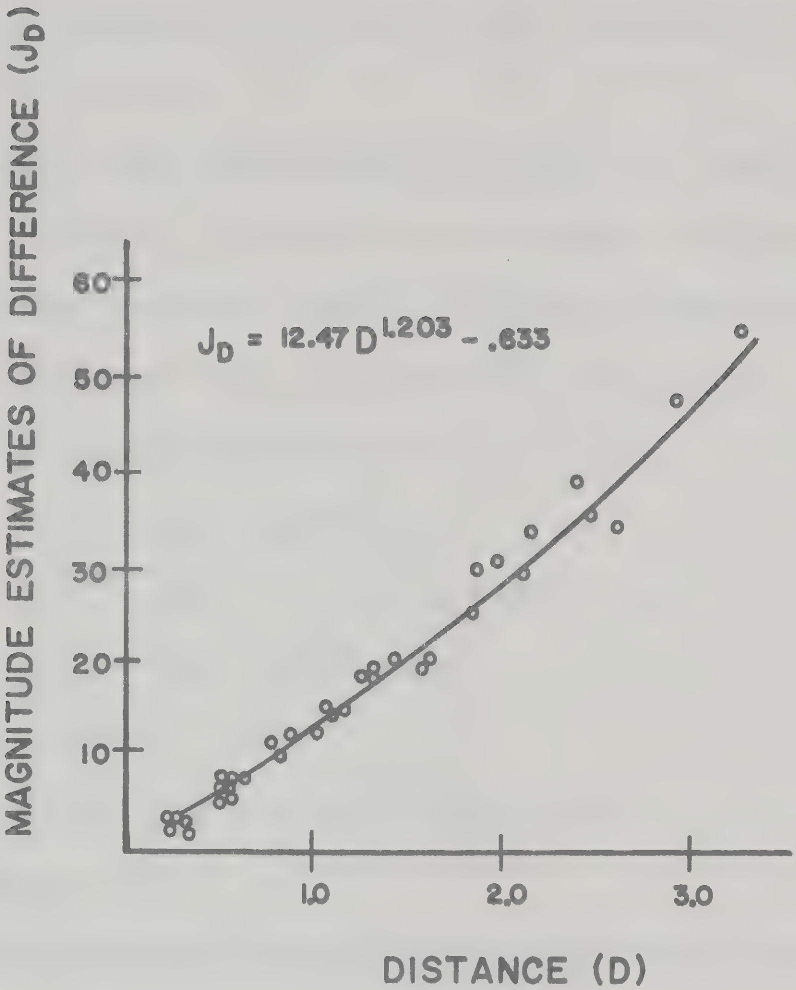

Figure 10. Magnitude estimations of difference as a function of interpoint distance (D) obtained from nonmetric scaling analysis.

against interpoint distance on the dimensions from the nonmetric analyses.

The output transformation in the two stage model is a power function. Therefore if the model is appropriate, difference judgments should be a power function of interpoint distance from the nonmetric analyses and the two stage model of Equation 6 should fit the data. To recapitulate:

$$V_i = a\emptyset_i{}^k - b, \tag{9R}$$ [1]

$$D_{ij} = V_i - V_j = a(\emptyset_i{}^k - \emptyset_j{}^k), \tag{12}$$

$$J_d = a_1\ D^m + b_1, \tag{13}$$

and

$$J_d = a_2\ (\emptyset_i{}^k - \emptyset_j{}^k)^m + b_1, \tag{6R}$$

where $a_2 = a_1\ a^m$.

The Category Rating of Difference data from the present study were described quite adequately by a power function

$$J_d = 2.28\ D_{ij}{}^{.87} + .547,\ r^2 = .983, \tag{14}$$

where D_{ij} refers to interpoint distance between stimuli $\underline{i}$ and $\underline{j}$, and $\underline{r^2}$ indicates proportion of dependent variable variance accounted for by the function. For the range of distance involved Equation 14 is virtually indistinguishable from a linear function. The exponent is close to 1.0.

For the Magnitude Estimation of Difference data a

---

1. When numbering equations, R is used whenever an equation is repeated.

power function

$$J_d = 12.4(D_{ij})^{1.20} - .63, \; r^2 = .984, \qquad (15)$$

was found to give a satisfactory fit.

The results in Table 2 and Equations 14 and 15 provide estimates of the parameters a, a_1, k, b and m of Equation 6. These estimates for the magnitude estimation and category rating of difference data were

$$J_{md} = 94.9(\emptyset_i^{.43} - \emptyset_j^{.43})^{1.20} + .633,$$

and

$$J_{cd} = 9.359\,(\emptyset_i^{.49} - \emptyset_j^{.49})^{.87} + .547.$$

Least squares solutions for Equation 6, obtained directly from the data, were for the magnitude estimations of difference

$$J_{md} = 97.09\,(\emptyset_i^{.44} - \emptyset_j^{.44})^{1.24} + .847 \qquad (16)$$

and for the category ratings of difference

$$J_{cd} = 9.51\,(\emptyset_i^{.49} - \emptyset_j^{.49})^{.905} + .656. \qquad (17)$$

Equations 16 and 17 accounted for 97 and 99 percent of the original dependent variable variance.

The input and output exponents of Equations 16 and 17 fail to adequately predict the lightness exponents of Figures 1 and 2. Multiplying .44 by 1.24 yields .55, which is not the value, 1.73, found in Figure 2. Similarly, .49 x .905 = .445; whereas n of Figure 1 was .64.

Similarity Functions

There are many possible models that could be developed to relate psychological similarities to distances and to judgments of these quantities. One model felt to best describe the present data will be presented here. Other possible models are described later.

One of the simplest ways to conceive of the relationship between the subjective similarities (S_{ij}) and subjective differences (D_{ij}) between stimuli is the notion that similarity is a complementary function of difference:

$$S_{ij} = G - D_{ij}. \qquad (18)$$

The constant G may be considered, at present, an arbitrary fitting constant. The implications of this constant are developed later. Assume that the output operations for similarities are like those employed in judgments of differences, that is,

$$J_{S_{ij}} = a\ S_{ij}{}^{m} + b. \qquad (19)$$

Substituting Equation 18 into Equation 19 yields

$$J_{S_{ij}} = a(G - D_{ij})^{m} + b, \qquad (20)$$

as the model relating judged similarities to interpoint distances.

Translation of Equation 20 into a model relating judgments to physical reflectances begins by replacing

D_{ij} with the right side of Equation 12:

$$J_{S_{ij}} = a[G - \alpha(\phi_i^{\,k} - \phi_j^{\,k})]^m + b.$$

Moving α outside the brackets yields

$$J_{S_{ij}} = a\alpha^m \left[\frac{G}{\alpha} - (\phi_i^{\,k} - \phi_j^{\,k})\right]^m + b.$$

If $G' = \frac{G}{\alpha}$ and $a' = a\alpha^m$ then

$$J_{S_{ij}} = a'[G' - (\phi_i^{\,k} - \phi_j^{\,k})]^m + b, \qquad (21)$$

relates judged similarities to reflectance value.

Least square solutions for the model of Equation 20 with the three sets of similarity data yielded

$$J_{cs} = 1.99(3.09 - D_{ij})^{1.06} + .942 \qquad (22)$$

for the category rating of similarity data,

$$J_{se} = 9.23\ (3.848 - D_{ij})^{1.805} + 2.53 \qquad (23)$$

for the similarity estimation, and for the magnitude estimations of similarity

$$J_{ms} = .088(3.8 - D_{ij})^{4.32} + 3.47. \qquad (24)$$

These functions are shown in Figures 11, 12 and 13. Equations 22 and 23 had r^2 values of .98 or greater. The magnitude estimation of similarity data could not be fit well. Residual variance for Equation 24 was 7 per cent of the original variance of the dependent measure.

A least squares fit of Equation 21 to the similarities data yielded

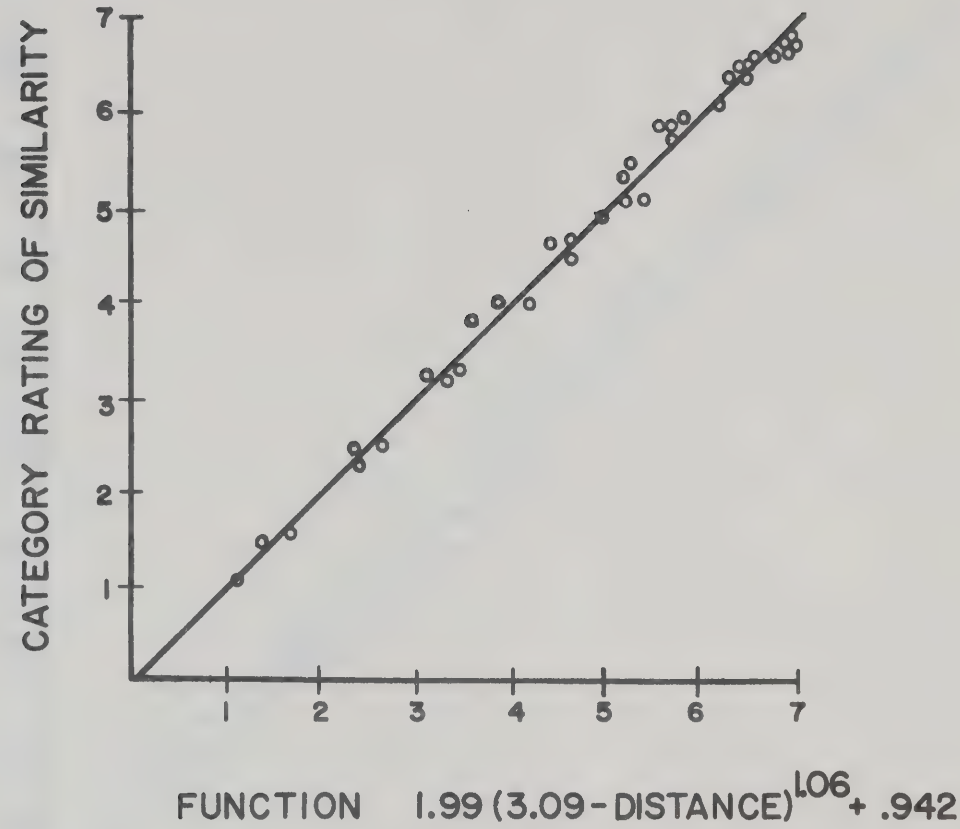

Figure 11. Category rating of similarity as a function of interpoint distance. Distance values are derived from nonmetric scaling analysis.

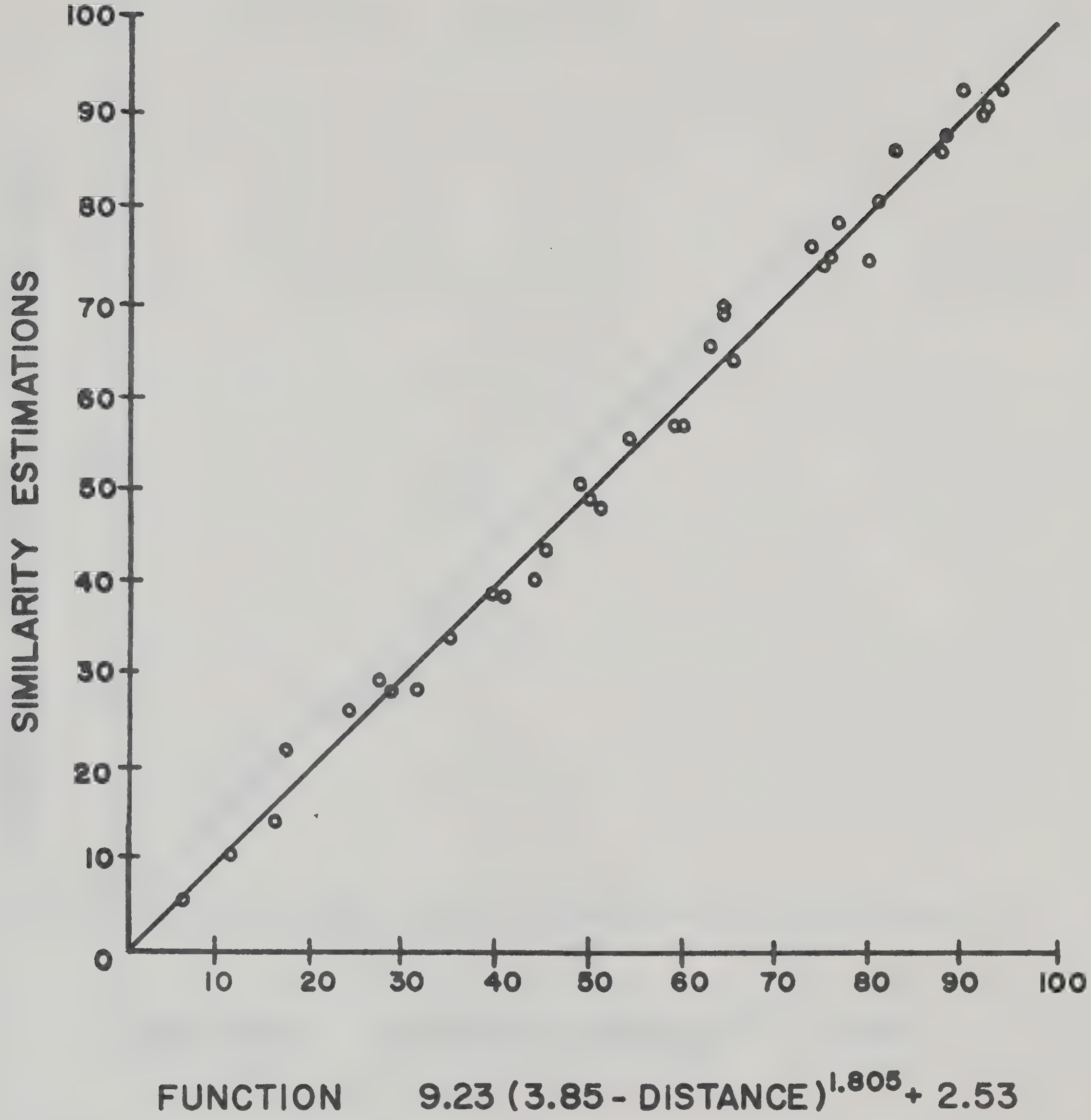

Figure 12. Similarity estimations as a function of interstimulus distance.

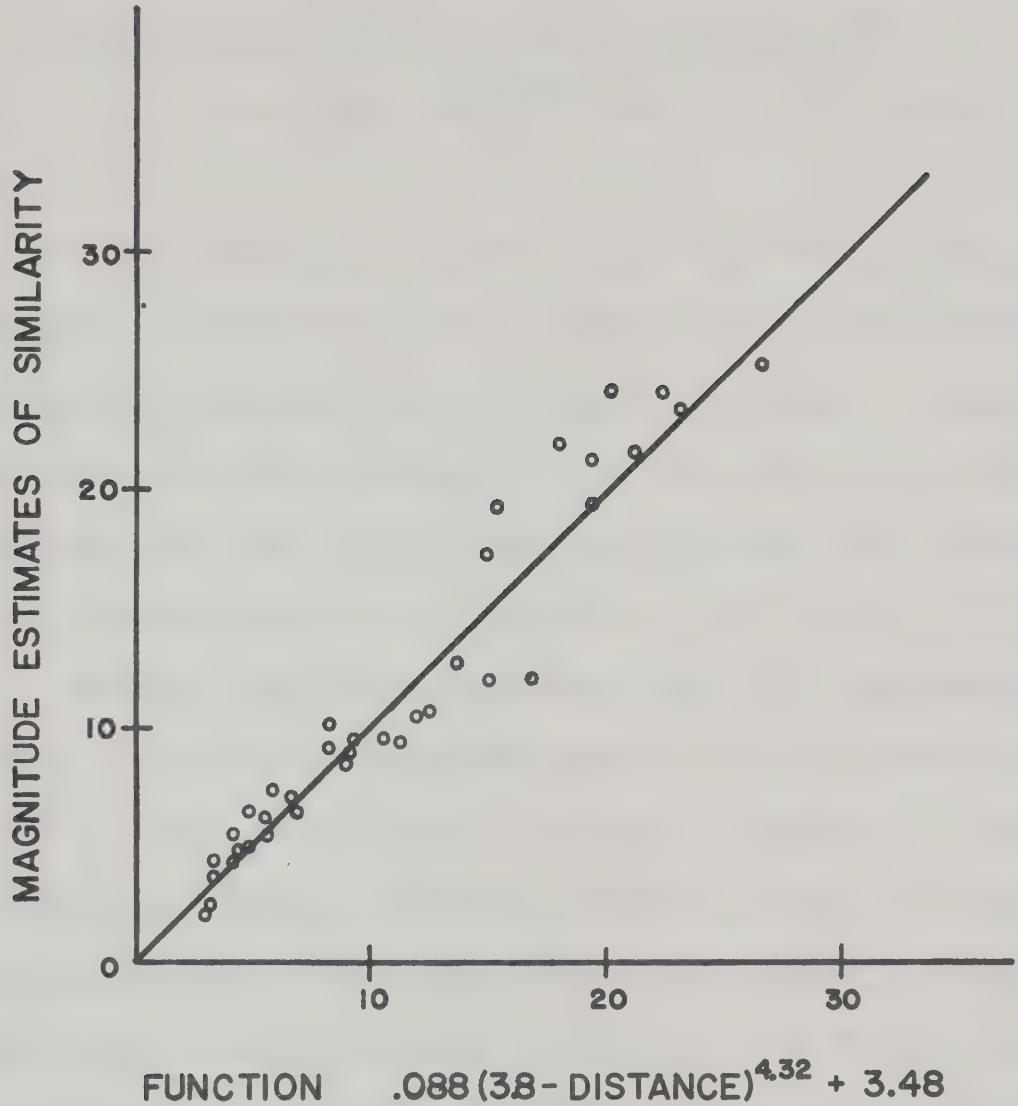

Figure 13. Magnitude estimations of similarity as a function of interpoint distance.

$$J_{cs} = 8.92\left[.721 - (\varnothing_i^{.857} - \varnothing_j^{.857})\right]^{1.088} + 1.0, \quad (25)$$

$$J_{se} = 110.3\left[.991 - (\varnothing_i^{.71} - \varnothing_j^{.71})\right]^{1.896} - 2.526, \quad (26)$$

and

$$J_{ms} = 24.35\left[1.0 - (\varnothing_i^{.797} - \varnothing_j^{.797})\right]^{3.88} + 2.759, \quad (27)$$

for the category rating of similarity, similarity estimation and magnitude estimation of similarity data sets. Once again the category rating and similarity estimation data were fit very well ($r^2 > .98$) while the magnitude estimation function was not as satisfactory ($r^2 = .89$).

Several questions were raised about the interrelationships of the various response procedures at the beginning of the study. Very briefly, magnitude estimates were not linear with category ratings. Category ratings of similarity appeared to be complementary to category ratings of difference. Also magnitude estimates of similarity appeared to be reciprocally related to magnitude estimates of difference. However consideration of the equations of the two stage model indicates that such a view may be too simple. Figures and equations relating the various response procedures are found in Appendix III.

DISCUSSION

Previous authors have discussed possible multistage processes to account for the commonly found inter-relations between scaling judgments and response procedures. Treisman (1964), for example, has shown that a two step judgment process involving a log input stage and an exponential output stage would account for many of the empirical findings from Stevens' Power Law.

The question is whether a coherent multistage model can be constructed that will clearly account for the types of relationships found in the present study.

The first requirement for a model is an input operation, that is, a psychophysical function relating Ss' sense impressions to physical measures. There are two possibilities in the literature: (1) Fechner's log law and (2) the notion that a subjective magnitude is a power function of physical magnitude. Evidence separating the two ideas should be gained from observing the relationship between the scale values from nonmetric scaling and the physical measure, reflectance. The present data support a power function; whereas Markley, Ayers, and Rule (1969) found evidence for the log function.

Additional assumptions are needed to convert a subjective magnitude into a judgmental response. The differences between various response procedures suggest a system wherein instructions serve either to alter parameter values of a single output function or change altogether the nature of the function.

There is some support in the literature for a multiple function system. It has frequently been reported that category ratings and magnitude estimations are logarithmically related to one another. (Although, the relationship usually shows less curvature than a log function.) Thus, a model such as that proposed by Treisman which has a log input function requires a linear output to account for category ratings and an exponential output to deal with magnitude estimations.

Attneave (1962) suggested that power functions serve both for input and output operations. Curtis, Attneave, and Harrington (1968) developed this notion quantitatively into a two stage model describing magnitude estimation judgments of weights and of differences in weights. Later evidence (Curtis, 1968; 1970) extended this model to category rating judgments and to other modalities. In this model the differences usually observed between

magnitude estimations and category ratings represent only a change in the value of the output function's exponent and not a qualitative change in function type.

A Model

A model similar to that of Attneave and co-workers is outlined in the equations presented below and can summarize most of the results obtained in the present study. Table 3 presents the basic equations and parameter estimates obtained from the data collected in this study.

For a unidimensional stimulus $\emptyset_i$, its subjective magnitude (lightness), V_i, is a power function of physical magnitude.

$$V_i = a_1\emptyset_i^{\,k} - b. \qquad (9R)$$

Empirical estimates of the values of the parameters of Equation 9 are found in Table 2. A combined estimate of k was .64. The category ratings of apparent magnitude (lightness judgments) and magnitude estimations represent a power transformation of V_i. That is, in general,

$$J = a_2(V_i + b)^m + b_2. \qquad (11R)$$

Estimates of m were 1.02 and 2.86 respectively for the category ratings and magnitude estimations of lightness.

The relationship between judgments (J) and physical magnitude ($\emptyset$) is also a power function. Combining

TABLE 3

SUMMARY OF RESULTS

	Difference Judgments					
	$V = a\emptyset^k + b$	$J_d = aD^m + b$		$J_d = a(\emptyset_i^k - \emptyset_j^k)^m + b$		
	k	m		k	m	
Magnitude Estimations	.43	1.20		.44	1.24	
Category Rating	.49	.87		.48	.90	
	Similarity Judgments					
	$V = a\emptyset^k + b$	$J_s = a(G - D)^m + b$		$J_s = a[G' - (\emptyset_i^k - \emptyset_j^k)]^m + b$		
	k	m	G	k	m	G'
Category Rating	.86	1.06	3.09	.857	1.09	.72
Similarity Estimation	.63	1.80	3.85	.710	1.90	.99
Magnitude Estimation	.67	4.32	3.80	.797	3.88	1.0
Combined Scales (Similarities and Differences)	.635					

Equation 9 and Equation 11 yields

$$J_i = a_2\, a_1{}^m \emptyset_i{}^{km} + b_2,$$

let

$$a_3 = a_2 a_1{}^m, \text{ and } n = km$$

then

$$J_i = a_3 \emptyset^n + b_2. \tag{10R}$$

The values of n observed for the present data were .64 and 1.73 respectively for category ratings and magnitude estimations of lightness. Agreement between $\hat{n}$ values predicted from k and m and observed n values was excellent.

Two variables that are both power functions of a third variable are also related to each other by a power function. Magnitude estimations and category ratings should be related to one another by a power function with an exponent of 1.71/.64 = 2.68. The least squares estimate of this exponent, shown in Figure 3, was 2.92. Over the range of values shown the difference between functions with exponents of 2.68 and 2.92 is trivial.

Differences. The application of these ideas to the difference data was detailed previously. To reiterate briefly:

A subjective difference, D_{ij}, is equal to the difference between the subjective magnitudes (V_i) of the stimuli.

$$D_{ij} = V_i - V_j = a(\emptyset_i{}^k - \emptyset_j{}^k). \tag{12R}$$

Judgments of differences are then

$$J_d = a_1 D^m + b, \tag{13R}$$

replace a_1 by a and then

$$J_d = a(\emptyset_i{}^k + \emptyset_j{}^k)^m + b_2. \tag{6R}$$

For magnitude estimations of difference the output operation was found to be a power function with an exponent of approximately 1.2. The estimates of output exponents for category ratings of differences were .87 from nonmetric analysis and .90 from least squares estimates of Equation 6.

The values of the input exponents for all the difference data were consistent with those obtained by nonmetric analyses in Table 2 and prior research (Curtis, 1968). The value of the output exponents (m) for magnitude estimates of difference in the present study were greater than 1.00 which was consistent with values found for other continua by Curtis et al. (1968). However, the values differed from the computation of m for the magnitude estimations of lightness data (Figure 8), also, the expectation that k x m from the difference data would give the value of the exponent n computed for the lightness scales was not fulfilled. Curtis (1968) has reported the same inconsistency for Torgerson's (1960) data from judgments of the differences of greys.

The contradiction has not been found for other continua, e.g., brightness, heaviness and circle size. The input exponents (.43 through .49) reported here for differences do compare well with the input values (.41 and .37) calculated by Curtis (1968) for the Torgerson data. The failure to predict the lightness exponent could be a function of problems in determining the lightness exponent itself or in the value of the output exponent. Published estimates of lightness exponents vary from 2.36 (Torgerson 1960) to .45 Mashhour & Hosman 1968). There is no consensus about one exponent for human lightness judgments. Also, Curtis (1968) has reported that while estimates of the input exponent (k) are fairly stable the output exponent (m) appears to be highly variable and effected by a variety of situations and conditions, only some of which are under experimental control.

Similarities. The similarity model used in analyses of the present data was contained in Equation 18,

$$S_{ij} = G - D_{ij}, \tag{18R}$$

It was then assumed that overt judgments of similarity are a power function of S_{ij}.

$$J_s = aS^m + b, \tag{19R}$$

and

$$J_s = a(G - D_{ij})^m + b. \tag{20R}$$

Substituting physical measures for D_{ij} and rearranging coefficients leads to Equation 21.

$$J_s = a'\left[G' - (\emptyset_i{}^k - \emptyset_j{}^k)\right]^m + b. \tag{21R}$$

Within the three sets of similarity data, the computed values of k, the input exponent, were consistent with one another and with the values obtained from the nonmetric analyses of Table 2. The input exponents from the similarity data were somewhat higher than those computed for the difference data. The values of the output exponents, with the exception of that from the magnitude estimation of similarity, were consistent with the results of Curtis and the values expected of the model.

The output exponents computed for the Magnitude Estimation of Similarity data were much higher than expected. The data are perhaps too unstable to permit really good fits to be found for various functions. None of several different functions that were computed relating the magnitude estimation of similarity data to any of several possible independent variables was able to account for more than 93 percent of the unweighted dependent variable variance. For all other judgmental tasks (e.g., category rating of differences), r^2 values were routinely

around .98 percent or greater.

Other Models

With respect to the similarities data, there are several other models that could be developed. Several of these will be described below and related to parts of the present data.

Exponential Output Model. Three assumptions are made: (1) There is only one subjective relationship between stimuli and the same process is involved in perceived similarity and perceived difference. (2) The response processes operating in similarity judgments are not related to those operating upon impressions of single stimuli and differences between stimuli. (3) Similarity judgments are exponential transformations of distance.

The latter assumption found empirical support in the paper by Markley, Ayers, and Rule (1969). Their data was described by a function

$$J_{S_{ij}} = ae^{-bDij} \qquad (28)$$

where D_{ij} was a Kruskal distance between the ith and jth stimuli, a and b were fitted constants and $J_{S_{ij}}$ was similarity estimation response. Least squares solutions for Equation 28 applied to the present data resulted in

$$J_{CS_{ij}} = 8.06\ e^{-.47D_{ij}}, \quad r^2 = .974, \tag{29}$$

$$J_{SE_{ij}} = 115.0\ e^{-.63D_{ij}}, \quad r^2 = .980, \tag{30}$$

and

$$J_{MS_{ij}} = 31.8\ e^{-1.0D_{ij}}. \quad r^2 = .918, \tag{31}$$

for category rating of similarity, similarity estimation, and magnitude estimates of similarity, respectively.

If differences are related to $\emptyset$ by Equation 12,

$$D_{ij} = a_1\ (\emptyset_i{}^k - \emptyset_j{}^k), \tag{12R}$$

then combining equations 28 and 12 gives

$$J_{S_{ij}} = ae^{-B(\emptyset_i{}^k - \emptyset_j{}^k)} \tag{32}$$

with $B = a_1 b$. Least squares solutions of Equation 32 for the present similarities data yielded:

$$J_{CS} = 8.29\ e^{-2.01\ (\emptyset_i{}^{.819} - \emptyset_j{}^{.819})}, \ r^2 = .96, \tag{33}$$

$$J_{SE} = 24.75\ e^{-3.28(\emptyset_i{}^{.79} - \emptyset_j{}^{.79})}, \quad r^2 = .98. \tag{34}$$

$$J_{MES} = 137.7\ e^{-3.75(\emptyset_i{}^{.7} - \emptyset_j{}^{.7})}, \quad r^2 = .89. \tag{35}$$

Again, the magnitude estimates of similarity were not fitted in a completely satisfactory manner.

With the present data this exponential output model is not distinguishable from the two stage model presented earlier. Conceptually, the exponential model is not as parsimonious as the two stage model as it requires special processes to deal specifically with similarity judgments. No exponential output function could be found to adequately

describe the difference data.

Reciprocal Model. Another possible way to relate similarities to differences (or distance) is to view similarity as the reciprocal of distance

$$S_{ij} = \frac{1}{D_{ij}} . \tag{36}$$

If the response processes suggested by the two stage model analyses of the difference data operate also on similarities, then

$$J_s = a\ S_{ij}^{m} + b$$

$$= a \left[\frac{1}{D_{ij}}\right]^{m} + b$$

$$= a\ D_{ij}^{-m} + b. \tag{37}$$

A least squares solution of Equation 37 for the magnitude estimates of similarity resulted in

$$J_{S_{ij}} = 1.35.6\ (D_{ij})^{-.069} - 124, \quad r^2 = .904. \tag{38}$$

Equation 39 is a two parameter function obtained via the reduction technique (Lewis, 1960).

$$J_{S_{ij}} = 116.6\ (D_{ij})^{-1.078}, \quad r^2 = .897. \tag{39}$$

A similar development for the similarity estimation and category rating of similarity data failed when it was not possible to compute a function with a negative exponent. The computer program always iterated to a function with positive exponent and negative coefficient. Clearly, the

reciprocal model was not the best available for these data. Due to the poor fit of the distance functions, no attempt was made to replace distance values with physical values.

Simple Ratio Model. A third, alternative, similarity model is one that avoids use of differences or distances. The model may take two forms

$$S_{ij} = \frac{\emptyset_i}{\emptyset_j}, \tag{40}$$

where $\emptyset_i$ and $\emptyset_j$ are two physical measures of stimuli; or

$$S_{ij} = \frac{V_i}{V_j} = \frac{{\emptyset_i}^k}{{\emptyset_j}^k} \quad . \tag{41}$$

That is, a similarity is the ratio of either the physical intensities or subjective magnitudes of stimuli. The present data did not fit this model. For the magnitude estimates of similarity it was found that the least squares fitted function of the form

$$J_S = a\left(\frac{\emptyset_i}{\emptyset_j}\right)^m + b \tag{42}$$

accounted for 85 percent of the dependent variable variance. Like functions for the other similarity data sets were even less adequate.

Ekman's Model. Ekman and co-workers (cf., Ekman & Waern, 1959) have proposed a model of similarity relating judgments of similarity ($J_{S_{ij}}$) to independently obtained

judgments of the apparent magnitude (J_i) of single stimuli. The proposed model was

$$J_{S_{ij}} = \frac{2\ J_i}{J_i + J_j}, \qquad i < j. \tag{43}$$

The model offers little by way of direct application for the present study since it is a relationship between two response variables and contains no direct implications for internal psychological processes.

As a matter of interest, the Ekman model was fit using the similarity estimation (SE) and magnitude estimation of lightness (J_i) results. A linear relationship between predicted and observed similarity judgments was found. If

$$\hat{S}_{ij} = \left[\frac{2\ J_i}{J_i + J_j}\right] \left[\frac{100}{1}\right], \text{ where } i < j,$$

it was found that the function

$$SE = .996\ (\hat{S}) + 6.55 \tag{44}$$

accounted for 91 percent of the variance ($r = .954$) of the similarity judgments. The fit to the similarity estimates was not as good as that found for two stage model equations for these data. However, these data fit the model in Equation 43 better than any reported previously from outside Ekman's Stockholm laboratory. For example,

unreported analyses of the Markley *et al*. (1969) study found a correlation between predicted and observed similarity values of only .78.

Nature of G

Little has been said up to now about the possible psychological interpretation of the arbitrary constant, G, introduced in the similarity model of Equation 18. One possible interpretation would relate this parameter to the maximum difference expected by the *S* in the context of the experiment. That is,

$$G = D_{max} = \Psi_{max} - \Psi_{min} = a(\phi^k_{max} - \phi^k_{min}) \quad (45)$$

where Ψ is a subjective magnitude and ϕ a physical magnitude, repeating Equation 18

$$S_{ij} = G - D_{ij}, \quad (18R)$$

or

$$S_{ij} = (\Psi_{max} - \Psi_{min}) - (\Psi_i - \Psi_j), \quad (46)$$

an equivalent form, except for unit of measurement, is

$$S_{ij} = \frac{(\Psi_{max} - \Psi_{min}) - (\Psi_i - \Psi_j)}{(\Psi_{max} - \Psi_{min})}. \quad (47)$$

Equation 47, expressed in subjective units, is the model for similarity proposed by Carmichael *et al*. (1965). For all analyses reported here the denominator of Equation 47 can be thought of as being incorporated into a multiplicative

coefficient, e.g., the parameter a' in Equation 21.

Numerical estimates of G reported above were 3.09, 3.85, and 3.8 for the distance functions, and corresponding estimates of G' were .72, .99, and 1.0 for functions employing reflectance as an independent variable. The values of G are each nearly the same as the distance between the extreme stimuli on the single dimension found in the nonmetric analyses. The values for G' of around 1.0 are reasonable if we let Ψ_{min} be zero and Ψ_{max} be ϕ^k where ϕ is a reflectance of 1.0. The various coefficients involved with Ψ and ϕ values are again carried into the first coefficient a_1 of Equation 20.

Starting with

$$J_{S_{ij}} = a_1(G - D_{ij})^m + b, \tag{20R}$$

then substituting Equations 12 and 45 yields

$$J_{S_{ij}} = a_1(a\phi^k{}_{max} - a\phi^k{}_{min} - a\phi_i{}^k + a\phi_j{}^k)^m + b. \tag{48}$$

If $a\phi^k{}_{min}$ is subjectively of zero magnitude then it falls out of Equation 48. Also the quantity, a^m, can be moved outside the parentheses. For given constant lighting conditions, the brightest (and lightest) possible grey will have a reflectance of 1.0 or 100 percent. Therefore, $\phi^k{}_{max}$ is equal to 1.0, and

$$J_{S_{ij}} = a_1a^m(1.0 - \phi_i{}^k + \phi_j{}^k)^m + b. \tag{49}$$

Let $a_1 a^m = a'$ and Equation 49 is Equation 21 with G' equal to 1.0. Restricting G' to 1.0 and solving Equation 21 for the category data did not significantly alter the other parameter estimates. The residual variances were equal out to the fourth decimal place.

Without the simplifying assumptions used above the relationship between G and G' can be developed in another way. Equations 20 and 21 can be rearranged to yield

$$(J_s - b)^{1/m} = a^{1/m}G - a^{1/m}D$$

and

$$(J_s - b)^{1/m} = a'^{1/m}G' - a'^{1/m}(\phi_i^{\,k} - \phi_j^{\,k}).$$

From these equations it can be shown that

$$G' = \frac{a^{1/m}}{a'^{1/m}}\, G \qquad (50)$$

where $a^{1/m}$ is taken from Equation 20 and $a'^{1/m}$ is from Equation 21. Applying Equation 50 to the category rating of similarity data produces

$$\hat{G}' = \left[\frac{1.99^{1/1.06}}{8.92^{1/1.088}}\right](3.09) = .79$$

For the similarity estimation and magnitude estimation of similarity results estimates of G' were 1.10 and .95 respectively. If the two available estimates of m for each data set are pooled then for the category rating of

similarity data

$$\hat{G}' = (3.09)\left[\frac{1.99}{8.92}\right]^{2/(1.06 + 1.088)} = .77$$

Corresponding estimates for the similarity estimations and magnitude estimations are 1.01 and .97.

The model advocated above coupled an input (sensory) power function with an output (response) power function. The output exponent used by Ss varied with the E's instructions. Subjective similarities and differences were assumed to be simple transformations of the perceptual magnitude resulting from the input function. These occur before application of the output function. A particular output operation is thought to function in the same manner for any of the perceptual variables, be they single magnitudes, differences, or similarities, supplied by the sensory system's input operations.

It is worth noting that all estimates of the input exponent were clearly less than 1.0. They range from .86 to .41. Looking at the scale values plotted in Figure 6 it can be seen that if the five power functions were superimposed upon one another there would be little real difference between them. The value of the exponent is probably most strongly influenced by minor differences in

location of the extreme upper and lower end points. Although the input exponents from similarity conditions were all higher than those from difference conditions, there are not enough data yet available to consider the differences in values to be reliable.

Another important feature of these data was the values of the estimated output exponents (m) for the several category rating conditions. The values ranged from .87 to 1.08. All were near 1.0. Curtis (1968.1970) also reports m values from category ratings around 1.0. (Curtis also suggested that there were systematic deviations from a power function for his category data.) The implication is that category rating judgments may be linear transforms of subjective magnitudes. In fact, linear output functions fit to the present category data were indistinguishable from the power functions. Contrary to the argument made originally by Stevens in the development of his Power Law, it appears that magnitude estimations are a more serious distortion of the subjective scale than are category ratings.

Output exponents for the magnitude estimation conditions were varied. The one consistency was that the exponents were all greater than 1.0 and greater than output exponents from category rating conditions.

The fit of the model to the data was fairly good. Only the Magnitude Estimation of Similarity results were not described well. Certain quantitative relationships between observed parameters--in particular, the two stage model prediction that the product of the magnitude estimation of difference input (k) and output (m) exponents would equal the exponent obtained for magnitude estimation of lightness of greys--were not sustained. However, previous successful studies of the model have involved within Ss designs. The present study was a between Ss design. For that reason alone, the fact that the forms of the various functions are qualitatively appropriate is significant in light of the large individual differences in strategies commonly found in studies of human judgment.

Differences in strategies may also account for the differences between the present results and those of Markley et al. (1969). The ways in which the stimuli were similar were not specified directly by the instructions. (Subjects do tend to interpret category rating instructions in a more consistent manner than they interpret magnitude estimation instructions.) It may have been the case then that the Ss chose to employ different

similarity models in the present study than were used in the previous study. Some combinations of stimuli and response procedures may set the Ss to emphasize ratios of stimuli rather than differences in stimuli. This could produce the variation in input exponents reported here and the conflicting functions reported by Markley et al. (1969).

One result of the analyses was the power functions relating the category ratings to both physical reflectance and to the magnitude estimation responses. Previous findings have usually indicated that log functions relate these variables (an exception was Curtis, 1970). Logarithmic functions and power functions with fractional exponents are quite similar over the range of values usually studied in psychophysics. Perhaps the previous findings could also be described by power functions. The log relationship between category ratings and magnitude estimations was originally an empirical approximation. Later theorists attempted to rationalize the relationship (e.g., Helm, Messick, & Tucker, 1961; Eisler, 1962). The majority of studies have, in fact, not supported the log function but merely reported a "concave downward" relationship between category and magnitude estimation scales.

These curves were then interpreted to be an approximation to a log function (e.g., Stevens & Galanter, 1957).

REFERENCES

Attneave, F., Dimensions of similarity. <u>American Journal of Psychology</u>, 1950, <u>63</u>, 516 - 556.

Attneave, F., Perception and related areas. In S. Koch (Ed.), <u>Psychology: A Study of a Science</u>, Vol. 4, New York: McGraw-Hill, 1962, 619 - 659.

Carmichael, J. W., Julius, R. W., & Martin, P. D. M. Relative similarities in one dimension. <u>Nature</u>, 1965, 208, 544 - 547.

Cliff, N. Multidimensional scaling and cognition: I. Relations among judgments obtained under varying conditions. Technical Report No. 1, November, 1965, University of Southern California, Contract Nonr-228-32, Office of Naval Research.

Curtis, D. W. A two stage model of quantitative psychophysical judgment. Paper presented at the meeting of the Psychonomic Society, Saint Louis, Mo., October, 1968.

Curtis, D. W. Magnitude estimations and category judgments of brightness and brightness intervals. <u>Journal of Experimental Psychology</u>, 1970, <u>83</u>, 201 - 208.

Curtis, D. W., Attneave, F., & Harrington, T. A test of a two stage model of magnitude judgment. <u>Perception & Psychophysics</u>, 1968, <u>3</u>, 25 - 31.

Curtis, D. W. & Fox, B. E. Direct quantitative jugments of sums and a two stage model for psychophysical judgments. <u>Perception & Psychophysics</u>, 1969, <u>5</u>, 89 - 93.

Eisler, H. On the problem of category scales in psychophysics. <u>Scandinavian Journal of Psychology</u>, 1962, <u>3</u>, 81 - 87.

Eisler, H. & Ekman, G. A mechanism of subjective similarity. *Acta Psychol., Amst.*, 1959, *14*, 1 - 10.

Ekman, G. A direct method for multidimensional ratio scaling. *Psychometrika*, 1963, *28*, 33 - 41.

Ekman, G., Goude, G. & Waern, Y. Subjective similarity in two perceptual continua. *Journal of Experimental Psychology*, 1961, 61, 222 - 227.

Ekman, G. & Sjoberg, L. Scaling. *Annual Review of Psychology*, 1965, *16*, 451 - 474.

Ekman, G. & Waern, Y. A second-order ratio scale. *Acta Psychol., Amst.*, 1959, *16*, 78 - 80.

Fagot, R. F. & Stewart, M. Tests of product and additive scaling axioms. *Perception & Psychophysics*, 1969, *5*, 117 - 123.

Galanter, E. S. Contemporary psychophysics. In R. Brown, E. Galanter, E. Hess, & G. Mandler. *New Directions in Psychology*, Vol. I, New York; Holt; Rinehart, & Winston, 1962.

Garner, W. R., Hake, H. & Eriksen, C. Operationism and the concept of perception. *Psychological Review*, 1956, *63*, 149 - 159.

Hake, H. & Rodwan, A. Perception and recognition. In Sidowski, J. (Ed.), *Experimental Methods and Instrumentation in Psychology*. New York: McGraw-Hill, 1966.

Helm, C. E. Multidimensional ratio scaling analysis of perceived color relations. *Journal of the Optical Society of America*, 1964, *54*, 256 - 262.

Helm, C. E., Messick, S. & Tucker, L. R. Psychological models for relating discrimination and magnitude estimation scales. *Psychological Review*, 1961, *68*, 167 - 177.

Indow, T. & Uchizono, T. Multidimensional mapping of munsell colors varying in hue and chroma. Journal of Experimental Psychology, 1960, 59, 321 - 329.

Indow, T. & Kanazawa, K. Multidimensional mapping of munsell colors varying in hue, chroma, and value. Journal of Experimental Psychology, 1960, 59, 330 - 336.

Kruskal, J. B. Nonmetric multidimensional scaling: A numerical method. Psychometrika, 1964, 29, 115 - 129.

Landahl, H. D. Neural mechanisms for the concepts of difference and similarity. Bulletin of Mathematical Biophysics, 1945, 7, 83 - 88.

Markley, R. P. Subject effects in cross modality matching. Unpublished M.Sc. Thesis, University of Alberta, 1965.

Markley, R., Ayers, D. & Rule, S. Similarity judgments of line length. Perception & Psychophysics, 1969, 6, 58 - 60.

Mashhour, M. & Hosman, J. On the new "psychophysical law": A validation study. Perception & Psychophysics, 1968, 3, 367 - 375.

McGee, V. The multidimensional analysis of 'elastic' distances. British Journal of Mathematical and Statistical Psychology, 1966, 19, 181 - 196.

Melton, A. W. & Safier, D. E. Pairs of two syllable adjectives arranged in order of rated similarity. Unpublished study listed in E. R. Hilgard, Methods and procedures in the study of learning. In S. S. Stevens (Ed.), Handbook of Experimental Psychology, New York: Wiley, 1951.

Poulton, E. G. The new psychophysics: Six models for magnitude estimation. Psychological Bulletin, 1968, 69, 1 - 19.

Rule, S. J. Subject differences in exponents of psychophysical power functions. Perceptual and Motor Skills, 1966, 23, 1125 - 1126.

Rule, S. J. Subject differences in exponents for circle size and proportion. Perceptual and Motor Skills, 1968, 26, 520.

Rule, S. J. Subject differences in exponents from circle size, numerousness, and line length. Psychonomic Science, 1969, 15, 284 - 285.

Rule, S. and Markley, R. Subject differences in magnitude estimation and cross-modality matching. Perception & Psychophysics, (In press 1970).

Runquist, W. N. Verbal behavior. In Joseph B. Sidowski (Ed.), Experimental Methods and Instrumentation in Psychology. New York: McGraw-Hill, 1966, 487 - 540.

Shepard, R. N. The analysis of proximities: multidimensional scaling with an unknown distance function. I. Psychometrika, 1962, 27, 125 - 140.

Shepard, R. N. Attention and the metric structure of the stimulus space. Journal of Mathematical Psychology, 1964, 1, 54 - 87.

Shepard, R. N. Metric structures in ordinal data. Journal of Mathematical Psychology, 1966, 3, 287 - 315.

Stevens, S. S. On the psychophysical law. Psychological Review, 1957, 64, 153 - 181.

Stevens, S. S. & Galanter, E. Ratio scales and category scales for a dozen perceptual continua. Journal of Experimental Psychology, 1957, 54, 377 - 411.

Torgerson, W. S. Theory and Methods of Scaling. New York: J. Wiley, 1958.

Torgerson, W. S. Quantitative judgment scales. In H. Gulliksen & S. Messick (Ed.), Psychological Scaling: Theory and Applications. New York: Wiley, 1960, 211 pp.

Torgerson, W. S. Distances and ratios in psychophysical scaling. Acta Psychol., Amst., 1961, 19, 201 - 205.

Torgerson, W. S. Multidimensional scaling of similarity. Psychometrika, 1965, 30, 379 - 393.

Treisman, M. Sensory scaling and the psychophysical law. Quarterly Journal of Experimental Psychology, 1964, 16, 11 - 22.

Warren R. M. & Warren, R. P. A critique of S. S. Stevens' "new psychophysics." Perceptual and Motor Skills, 1963, 16, 797 - 810.

Zinnes, J. L. Scaling. Annual Review of Psychology, 1969, 20, 447 - 448.

APPENDIX I

STATISTICS SUMMARIZING RESPONSES TO PAIRS OF GREYS

MAGNITUDE ESTIMATION OF SIMILARITY (GEOMETRIC MEANS)

MUNSELL VALUE	9.5	9.0	8.5	8.0	7.5	7.0	6.0	5.0	3.5
9.5									
9.0	19.71								
8.5	12.73	24.56							
8.0	8.85	12.13	21.86						
7.5	7.03	9.64	19.04	23.54					
7.0	5.84	8.30	10.80	21.40	25.40				
6.0	5.83	7.73	8.96	10.63	12.15	24.47			
5.0	3.56	4.68	6.40	6.75	9.57	9.22	22.27		
3.5	2.36	2.63	4.60	4.83	5.14	6.32	10.22	17.541	

MAGNITUDE ESTIMATION OF DIFFERENCE (GEOMETRIC MEANS)

MUNSELL VALUE	9.5	9.0	8.5	8.0	7.5	7.0	6.0	5.0	3.5
9.5									
9.0	2.765								
8.5	7.384	3.37							
8.0	12.08	7.74	3.61						
7.5	15.32	10.39	6.48	3.27					
7.0	20.27	15.40	11.55	5.76	2.80				
6.0	30.54	20.04	18.01	15.61	11.27	5.07			
5.0	35.52	33.78	25.48	19.72	18.33	12.50	6.77		
3.5	54.94	47.86	33.99	39.56	29.78	29.98	18.88	11.07	

APPENDIX I (CONT.)

CATEGORY RATING OF DIFFERENCE
(MEDIANS)

MUNSELL VALUE	9.5	9.0	8.5	8.0	7.5	7.0	6.0	5.0	3.5
9.5									
9.0	1.11								
8.5	2.07	1.68							
8.0	2.86	2.23	1.42						
7.5	3.12	2.56	1.93	1.35					
7.0	3.64	3.14	2.42	1.83	1.21				
6.0	4.90	3.50	3.32	3.08	2.41	1.71			
5.0	5.78	5.50	4.83	3.83	3.19	2.87	1.77		
3.5	6.81	6.30	5.67	5.30	4.75	4.62	3.64	2.27	

CATEGORY RATING OF SIMILARITY
(MEDIANS)

MUNSELL VALUE	9.5	9.0	8.5	8.0	7.5	7.0	6.0	5.0	3.5
9.5									
9.0	6.76								
8.5	6.00	6.66							
8.0	4.92	5.90	6.71						
7.5	4.33	5.33	6.17	6.85					
7.0	4.04	4.50	5.50	6.50	6.88				
6.0	2.50	3.28	4.60	5.12	5.78	6.40			
5.0	1.64	2.28	3.21	4.00	4.67	5.50	6.58		
3.5	1.08	1.50	2.50	3.25	3.77	4.00	5.10	6.43	

APPENDIX I (CONT.)

SIMILARITY ESTIMATIONS
(ARITHMETIC MEANS)

MUNSELL VALUE	9.5	9.0	8.5	8.0	7.5	7.0	6.0	5.0	3.5
9.5									
9.0	93.21								
8.5	78.75	90.33							
8.0	57.21	76.29	88.38						
7.5	49.00	65.83	76.25	93.12					
7.0	38.38	47.71	70.25	81.96	91.29				
6.0	28.04	38.46	50.83	69.21	73.67	86.83			
5.0	14.60	26.46	28.08	43.00	55.62	63.50	87.33		
3.5	5.85	10.52	22.10	29.59	33.46	40.09	56.88	75.04	

APPENDIX II

SUMMARY OF SCALING OF LIGHTNESS AND DARKNESS

STIMULI		LIGHTNESS					DARKNESS				
		MAGNITUDE ESTIMATIONS			CATEGORY RATINGS		MAGNITUDE ESTIMATIONS			CATEGORY RATINGS	
MUNSELL VALUE	REFLECTANCE	gm	$\overline{X}$	sd	$\overline{X}$	sd	gm	$\overline{X}$	sd	$\overline{X}$	sd
3.5	.095	1.177	1.34	.667	1.00	0.0	43.089	48.92	25.72	6.94	.232
5.0	.200	2.98	3.66	2.04	2.20	.516	28.20	31.0	13.81	5.61	.687
6.0	.305	5.61	6.07	2.308	3.325	.797	20.44	21.32	6.61	4.58	.937
7.0	.43	8.11	8.39	2.14	4.05	.875	13.28	13.78	3.88	3.91	.840
7.5	.51	8.89	9.10	1.77	4.20	.822	10.94	11.21	2.88	3.61	.494
8.0	.58	10.92	11.42	4.35	5.02	.973	8.65	9.14	3.17	3.08	.500
8.5	.69	15.81	16.64	5.42	5.225	.800	6.78	7.18	2.31	2.64	.723
9.0	.80	25.375	28.82	14.85	6.20	.607	4.44	4.82	2.04	2.03	.736
9.5	.87	34.197	38.93	19.91	6.875	.334	2.26	2.46	1.10	1.14	.350

Each entry based on 4 responses from each of 7 _S_s (28 total responses)

APPENDIX III

Interrelations Between Response Measures

The magnitude estimations of similarities and of differences reported above were not linearly related to the category ratings of similarities and differences. The relationships are shown in Figures 14 and 15. The lines through the data were fitted by nonlinear least squares solutions to functions derived from the equations of the two stage model developed earlier.

The relationship between category ratings of difference and category ratings of similarity is presented in Figure 16. Figure 17 shows the two sets of magnitude estimations, and Figure 18 relates the magnitude estimates of similarity to the similarity estimations.

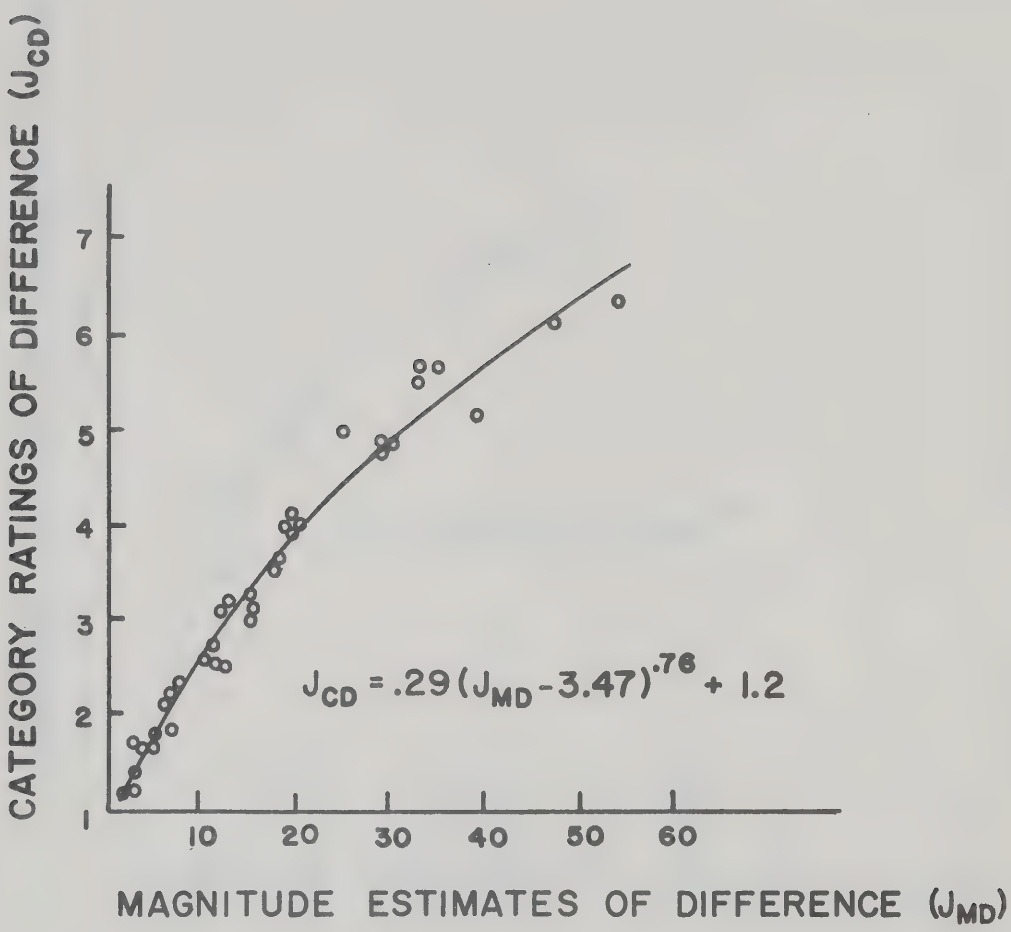

Figure 14. The relationship between category ratings and magnitude estimations of apparent difference.

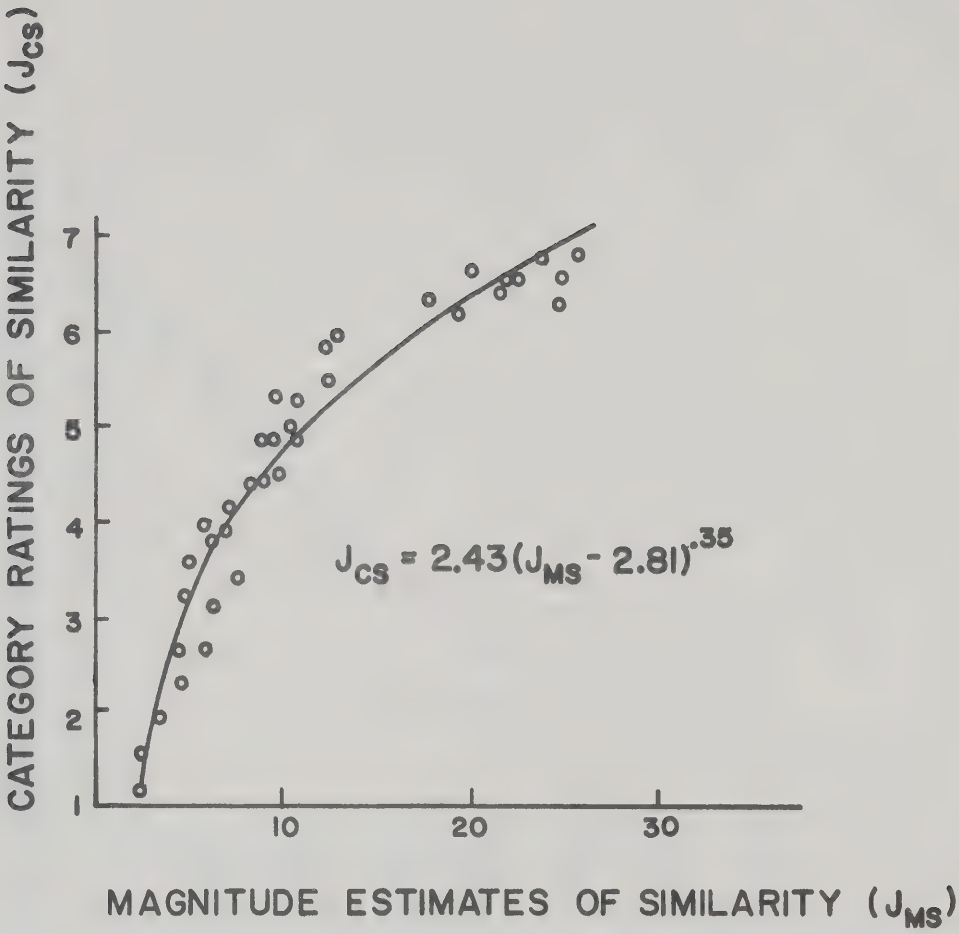

Figure 15. The relationship between category ratings and magnitude estimations of apparent similarity.

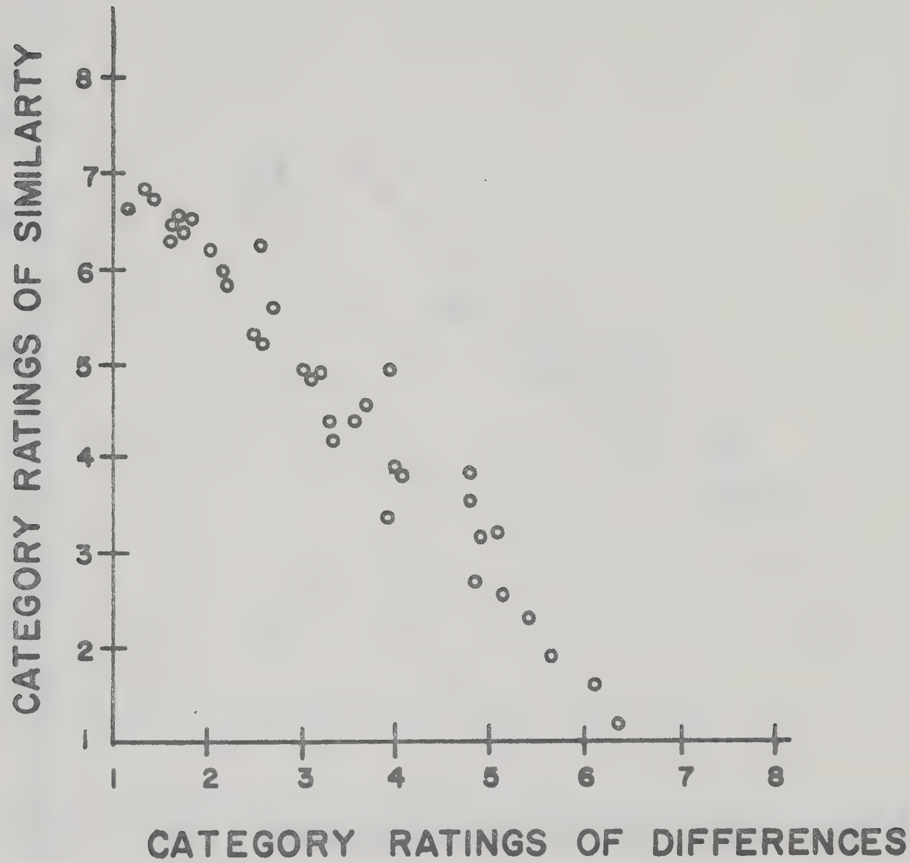

Figure 16. The relationship between similarities and differences of pairs of grey stimuli when responses are made using the category rating procedure.

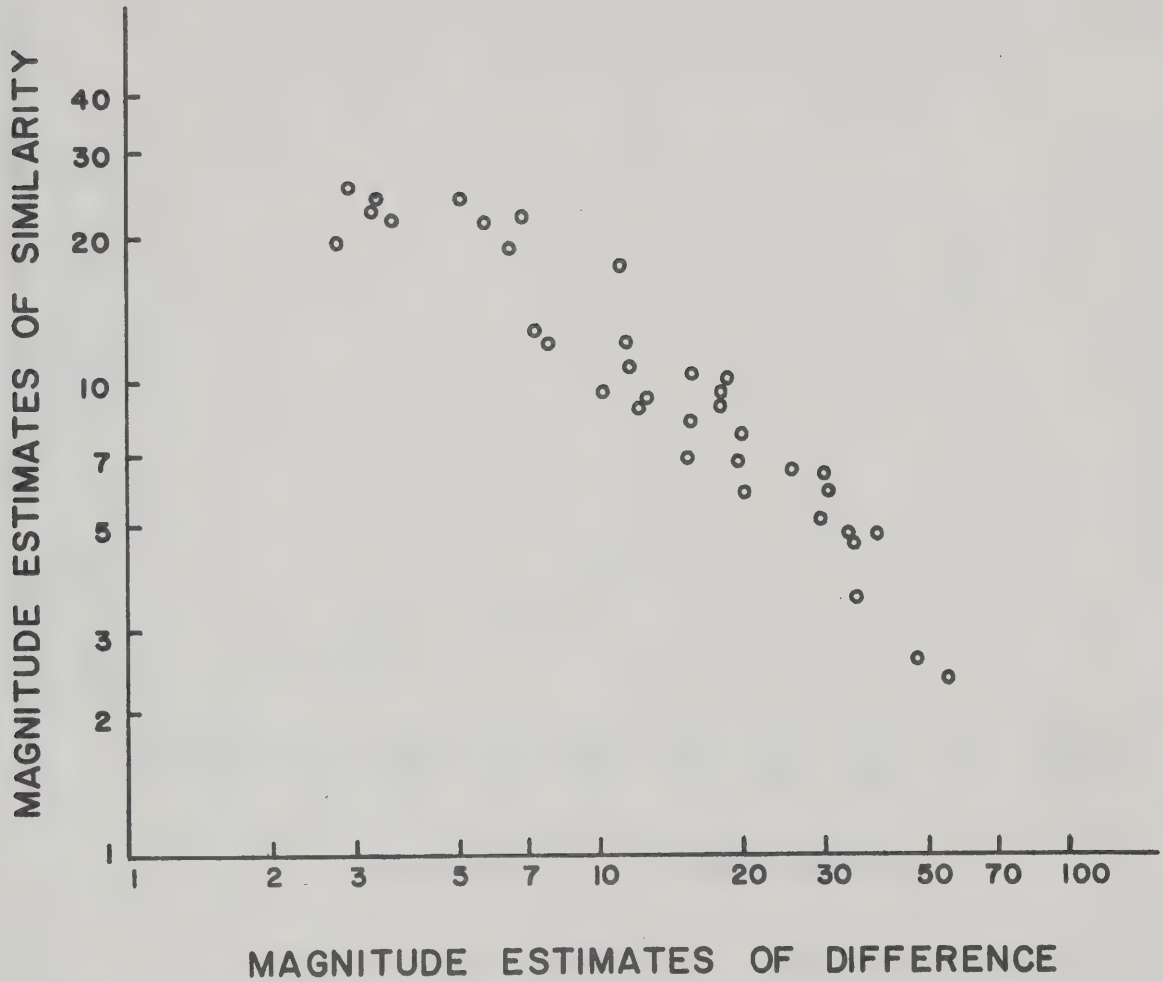

Figure 17. The relationship between magnitude estimates of similarity and magnitude estimates of difference.

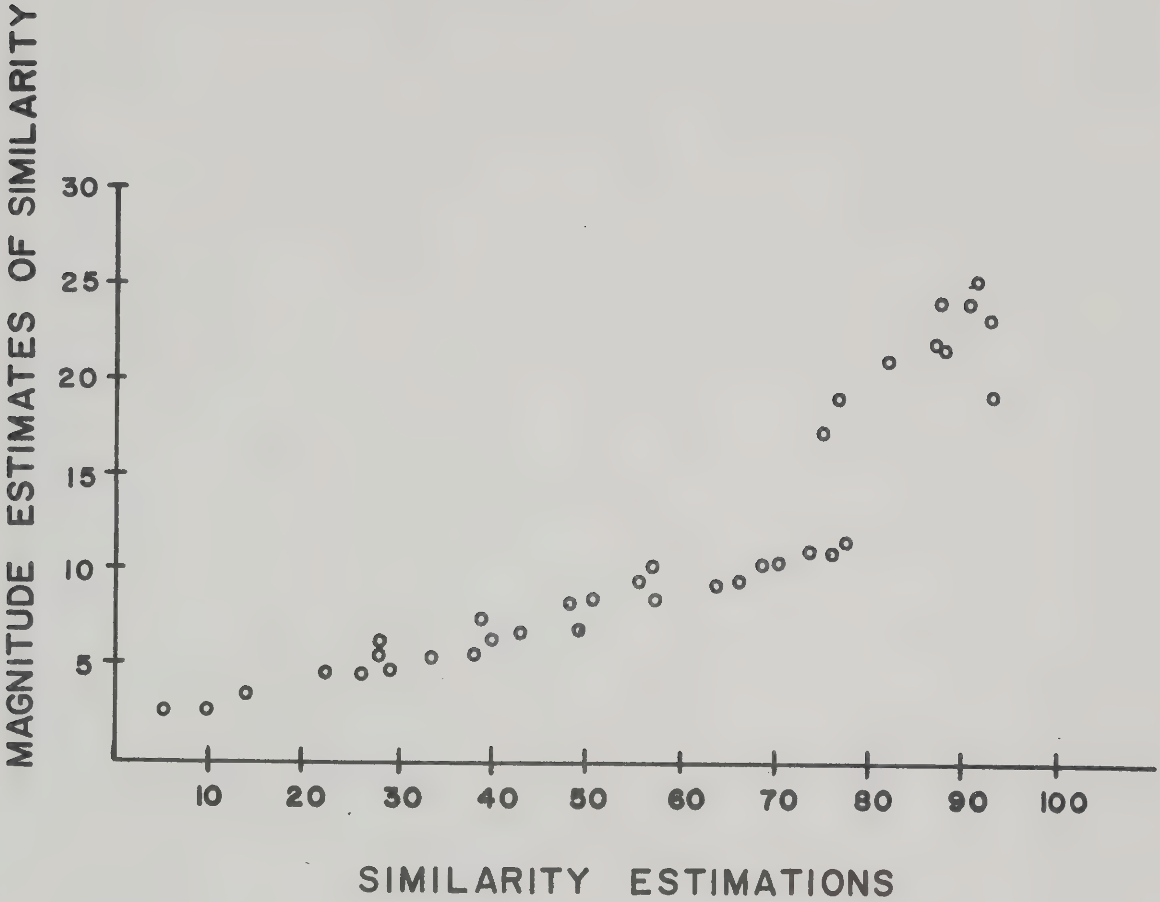

Figure 18. Magnitude estimations of similarity plotted as a function of similarity estimations.